...EPT.

THE AMERICAN MATHEMATICAL MONTHLY

THE OFFICIAL JOURNAL OF
THE MATHEMATICAL ASSOCIATION OF AMERICA, INC.

VOLUME 64 NUMBER 8

PART I

CONTENTS

The Practice of Mathematics . . . . . R. E. GASKELL 557
On Higher Ordered Differentiation . . . . DAVE PANDRES, JR. 566
On Certain Matrix Equations . . . . . T. J. PIGNANI 573
The Dirichlet Series Transformation . . . . . J. RAY HANNA 576
Mathematical Notes . RAYMOND REDHEFFER, PAUL J. MCCARTHY 582
Classroom Notes . . . . G. B. THOMAS, JR., R. M. CONKLING 586
Elementary Problems and Solutions 592
Advanced Problems and Solutions 596
Recent Publications 602
News and Notices 608
The Mathematical Association of America 612
New Members 612
April Meeting of the Metropolitan New York Section 614
April Meeting of the Missouri Section 616
May Meeting of the Allegheny Mountain Section 617
May Meeting of the Indiana Section 619
May Meeting of the Minnesota Section 622
May Meeting of the Southern California Section 623
May Meeting of the Upper New York State Section 625
May Meeting of the Wisconsin Section 627
June Meeting of the Pacific Northwest Section 629
Calendar of Future Meetings 630

OCTOBER 1957

PUBLIC LIBRARY

The AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

RALPH D. JAMES, *Editor*

ASSOCIATE EDITORS

RICHARD V. ANDREE	L. M. KELLY	EDITH R. SCHNECKENBURGER
HOWARD EVES	IVAN NIVEN	E. P. STARKE
G. E. HAY	C. O. OAKLEY	H. S. ZUCKERMAN
	J. M. H. OLMSTED	

EDITORIAL CORRESPONDENCE should be addressed to the Editor, R. D. JAMES, Department of Mathematics, University of British Columbia, Vancouver 8, Canada. Articles should be typewritten and double-spaced on 8½ by 11 paper. The greatest possible care should be taken in preparing the manuscript and authors should keep a complete copy.

ADVERTISING CORRESPONDENCE should be addressed to F. R. OLSON, University of Buffalo, Buffalo 14, N. Y.

NOTICE OF CHANGE OF ADDRESS by members of the Association as well as correspondence regarding subscriptions to the MONTHLY should be sent to the Secretary-Treasurer, H. M. GEHMAN, University of Buffalo, Buffalo 14, N. Y.

THIS IS THE OFFICIAL JOURNAL OF THE

MATHEMATICAL ASSOCIATION OF AMERICA, INC.

(Devoted to the Interests of Collegiate Mathematics)

OFFICERS OF THE ASSOCIATION

President, G. B. PRICE, University of Kansas
First Vice-President, R. V. CHURCHILL, University of Michigan
Second Vice-President, B. W. JONES, University of Colorado
Secretary-Treasurer, H. M. GEHMAN, University of Buffalo
Associate Secretary, EDITH R. SCHNECKENBURGER, University of Buffalo
Editor, R. D. JAMES, University of British Columbia

Additional Members of the Board of Governors: H. M. BACON, H. W. BRINKMANN, JEWELL H. BUSHEY, LAMBERTO CESARI, P. H. DAUS, W. L. DUREN, JR., H. P. EVANS, N. J. FINE, F. F. HELTON, M. S. HENDRICKSON, A. S. HOUSEHOLDER, W. N. HUFF, G. K. KALISCH, E. C. KIEFER, J. C. KNIPP, F. W. KOKOMOOR, W. G. LEAVITT, Z. L. LOFLIN, L. L. LOWENSTEIN, SAUNDERS MACLANE, E. J. MCSHANE, J. R. MAYOR, A. E. MEDER, JR., IVAN NIVEN, R. S. PARK, GEORGE PÓLYA, O. J. RAMLER, J. F. RANDOLPH, C. B. READ, C. R. SHERER, M. F. SMILEY, B. M. STEWART, G. B. THOMAS, JR., BERNARD VINOGRADE, C. R. WYLIE, JR., M. A. ZORN

Entered as second class matter at the post office at Menasha, Wis. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in Paragraph 4, Section 538, P. L. and R., authorized April 1, 1926.

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $5.00. For non-members the subscription price is $6.00.

PUBLISHED BY THE ASSOCIATION at Menasha, Wisconsin, and Buffalo, N. Y., during the months of January, February, March, April, May, June-July, August-September, October, November, December.

PRINTED IN THE UNITED STATES OF AMERICA

THE PRACTICE OF MATHEMATICS

R. E. GASKELL, Boeing Airplane Company

The place of the mathematician in our technology. Our society today benefits from the labors of several groups of dedicated individuals who are said to practice in certain fields. Physicians practice medicine, attorneys practice law, and we all know men who practice accounting, engineering or in other fields. We all recognize, too, that the most useful practicing physician may not be the one who knows the most about medicine, just as the most useful attorney may not be the one with the most profound knowledge of the law. Medical and legal training for these professions must of course be sound and thorough, yet it must be complemented with understanding and practical wisdom in many other fields.

Today we have an analogous situation in the field of mathematics. We have had professional mathematicians for some time, and now we find a growing number of practicing mathematicians. Because of the contributions of such pioneers as von Karman and Fry, the demand for mathematicians has increased tremendously, but to follow these pioneers faithfully, we must carefully control the product that we label "mathematician" if he is to meet the public professionally. To conform with the high standards set for practice in other professions, and for the practice of mathematics by those who were earliest in the field, we need mathematicians with more than a good solid grounding in mathematics. In addition, they need a sermonful of the attributes of the physician and attorney, plus considerable basic knowledge of engineering, physics and other sciences. As Dr. Fry explained some years ago, the mathematician in industry should be capable of doing a good engineering job if he turned his hand to it [1].* Today we find mathematicians assisting businessmen, production men, scientists, and many others in government and industry, so Fry's statement, though still true, should be broadly interpreted.

However, the consumer, whether he is an engineer, scientist, or businessman, will call upon the services of a mathematician only if and when he thinks he needs one. Too often he asks for help in an awkward situation that could have been avoided, or he asks for voluminous computations, with a "never mind what it's for or what method we're using." Of course, today's broader need for an understanding of mathematics means that businessmen, engineers and scientists must be far better trained in mathematics than was the custom fifteen years ago. Certainly more problems are solved, more sophisticated mathematics is used by these laymen—to borrow a much-borrowed word—than ever before. We must expect to see more mathematics of the do-it-yourself variety, and no one can really complain as long as the layman does not get beyond his depth. And since it is the layman who is first to judge the depth, the practice of mathematics is materially affected by what the public thinks mathematics and mathematicians are for.

* Numbers in brackets refer to the bibliography at the end of the article.

Fortunately, considerable mathematics does find its way to market, some of it under the name of Operations Research, but even so there is widespread ignorance as to what a mathematician is and what he can do. Not only is this ignorance widespread, but in a certain few individuals it is an intense, aggressive and contagious type of ignorance. Much credit is given to mathematicians in some quarters. Aerodynamicists will usually agree with von Karman when he points to the development of aerodynamic science as an "example of cooperation between 'men of mathematics' and creative engineers. Mathematical theories from the happy hunting grounds of pure mathematicians were found suitable to describe the airflow produced by aircraft with such excellent accuracy that they could be applied directly to airplane design" [2]. But we also read statements such as "There is great danger that the publicity given to these [computing] machines will develop an exaggerated idea of the value and place of mathematics in engineering design" [3].

We can only conclude that the position of the mathematician in our present day technology has not been uniformly established. Perhaps it can never be. Nevertheless, many people today are faced with the question, "Should the mathematical aspects of this study be carried out by an engineer or scientist with a flair for mathematics, or by a mathematician with a flair for the material under study?" In idler days this question would be interpreted as a bid for jurisdiction, but with work as abundant as it is, the question can only bear upon the efficiency of the utilization of our scientific manpower.

Popular notions about mathematics. It may be helpful to draw attention to the public attitude toward mathematics. Mathematicians must blame themselves for the widely held view that mathematics was all discovered years ago, and that it cannot possibly change—and therefore cannot grow. Our schoolchildren hear about Euclid and Pythagoras, our undergraduates hear about Apollonius, Descartes and Newton, and our mathematics majors may work up an acquaintance with the names of mathematicians who lived and worked in the nineteenth century—Cauchy, Fourier, Gauss, Laplace and others. Textbooks and classroom decorations hold the line in this respect, since custom seems to require that a mathematician spend a century in the grave before public display of his portrait is permitted.

It is common to find individuals who associate mathematical ability with winning at bridge (keeping score, anyway). Others place a mystic aura about the subject, as the executive who said that "Mathematics are quicker than the human eye" [4]. Some people fail to distinguish between accounting and mathematics; usually these folks just have bad consciences at income tax time and blame it on their poor arithmetic. The rash of cartoons depicting wheezing computers working unsuccessfully on income tax forms indicates that, in the public mind, computers have consciences, too.

Many other people, including most engineers and some scientists, equate mathematics with some of the more tangible tools of the mathematician, such

as formulas, graphs, charts, tables and computing machines. In this category we find those who place blind faith in the formula, those who delight in the various paradoxes concocted through use of extrapolation or sometimes through a sly division by zero.

The tremendous publicity given high-speed computing machines in the past five or six years has convinced many people that these machines are the very soul of mathematics. Well-known mathematicians have tripped on this one, too. One prominent mathematician was so completely overwhelmed that in reviewing a book on Laplace transformations he asked why the author bothered with Laplace transforms—why didn't he use a high-speed computer! It is quite true that the usefulness of mathematics has been increased enormously by the availability of high-speed computation, but it is a mistake to go further and assume that high-speed computers have replaced or will replace mathematics or mathematicians.

We mathematicians have our own ideas of what mathematics is. There are many definitions, of course. Mathematics has been called the queen of the sciences, and it also has been referred to as the handmaiden of the sciences—which led J. D. Williams, of RAND, to remark that while there is some doubt about the social position of mathematics, at least its sex is definitely known [5].

Alfred North Whitehead claimed for mathematics the distinction of being the most original creation of the human spirit. This led him to his definition, "Mathematics is thought moving in the sphere of complete abstraction from any particular instance of what it is talking about" [6]. And Bertrand Russell, Whitehead's colleague, figured he might as well agree with that, only he put it in plain English, defining mathematics as that science in which we do not know what we are talking about, nor whether what we say is true or not. This definition should enjoy almost universal acceptance. The mathematician accepts it because, in all seriousness, there is an element of truth in it, and by the time he explains its weaknesses, his listener will agree that it must be precisely correct. The non-mathematician accepts this definition readily—almost eagerly—because it explains so much of what he has heard and read.

In the face of these widely held and freely publicized notions, is it any wonder that a mathematician is so often called upon at the tail end of a decision and asked to "use this procedure to crank out a number on your little machine"?

The practice of mathematics. Many mathematicians in abstract branches of the subject must of necessity emphasize mathematical beauty and form. While it is true that their work finds applications, usually in another field of mathematics, we should compare their work to that of biological, pharmacological and medical researchers, who are among the leaders in the advancement of medical science, but who do not practice medicine. Louis Pasteur is the most notable example that comes to mind. We must agree, certainly, that this type of mathematician leads in the development of mathematical science.

Another type of mathematician works from the mathematics to the applica-

tion. He develops a mathematical method first, and then seeks instances in which he can apply it. He is an applied mathematician, but we should not call him a mathematical practitioner, at least not a *general* mathematical practitioner. He is more closely analogous to the medical specialist, who develops a technique and then stands ready to take over those cases in which his technique will prove beneficial.

Your true mathematical practitioner *does not seek the problem to fit his mathematics—he seeks that mathematics to which the problem can best be fitted*, with due regard for secondary effects, the quality of solution desired, time available for the solution, and the facilities available. Don't think for a moment that he ignores beauty and form, usually an elegant form goes hand in hand with efficiency. Nevertheless, his thought is *problem-centered*, rather than *method-centered*.

At this point let us boldly suggest that *the practice of mathematics involves the use of available and assimilable mathematical procedures to assist in a design or a decision*. The words *available* and *assimilable* are important here. A definitely superior method may exist, but time may be too short to develop it, the personnel assigned to carry out the work may not be trained enough in mathematics to use it safely, or the computing machinery necessary may not be accessible.

There seem to be five general activities in the practice of mathematics. For successful work all five must be practiced, though not necessarily by the same person or groups, and not necessarily in connection with every problem. Those who practice mathematics must

(1) *Recognize, understand and analyze the problem situation.** The man responsible for the design or the decision discusses the problem situation with a mathematical analyst. Together they pick out the fundamental points, and they strip away or ignore any irrelevant aspects. Compromise and negotiation usually feature these discussion sessions, which may be carried on intermittently for a considerable period of time—weeks, months, and sometimes for more than a year. The aim of these discussions is to arrive at a compromise which will be sufficiently realistic to satisfy the designer, yet practical for the mathematician to handle.

This activity, which may be called negotiation, almost always requires considerable ability in other sciences, in economics, business or in engineering.

(2) *Formulate the problem situation.* Formulation requires the ability to solve "word problems." Those who feel they have this ability to state problems in mathematical terms will be interested in Rufus Oldenburger's book "Mathematical Engineering Analysis" [7], where problems of this kind are emphasized. For example, Oldenburger poses the problem: "The bottom of a metal plate is

* The term *problem situation* is perhaps somewhat better than *problem* when it comes to describing the work of the practicing mathematician, though we shall use both terms interchangeably.

being uniformly heated by hot gases; the top is exposed to the atmosphere." That's the problem. The student takes over from there. Perhaps Oldenburger's masterpiece is this: "A pile of powdered coal is burning." That's it!

Continued negotiation is almost always necessary in formulating a problem for solution. Both parties must keep in mind the time, money, personnel and machinery available for its study. There is no point in formulating a problem precisely if it cannot be handled in finite time with the tools at hand, except that the precise formulation may persist as a challenge to develop useful tools for it at some future time.

(3) *Solve the problem stated* in mathematical form, negotiating modifications of the form where they are necessary to make the study practical. The solution of the problem is essentially the manipulation of its mathematical statement into an alternative form which can be interpreted, or which can be efficiently used in computation.

Let us consider a hypothetical illustration here. Suppose your client wants to design a pendulum which has a period of one second, and he is willing to neglect such things as air resistance, friction and movement at the point of support, elastic effects of the string, wire or bob, *etc.* Then your differential equation is nonlinear,

$$L\frac{d^2\theta}{dt^2} + g\sin\theta = 0.$$

If you have only a short time to get the solution, or if you do not have tables of elliptic integrals handy, you would negotiate with your client to replace $\sin\theta$ by θ, explaining that for small θ there isn't much error. Then

$$L\frac{d^2\theta}{dt^2} + g\theta = 0,$$

and this equation is easily solved:

$$\theta = A\sin t\sqrt{g/L} + B\cos t\sqrt{g/L}.$$

Note that this solution is interchangeable with the original differential equation from which it was obtained—hence the equivalence of the terms "solve" and "manipulate."

The period of this motion is determined from the period of the sine or cosine function, namely $T = 2\pi\sqrt{L/g}$. This is the formula usually given for the period of a pendulum.

But if your client will not allow restriction to small angles, and has the time and money for the extra effort, you would solve the differential equation as it stands, thus the period turns out to be

$$T = 4\sqrt{L/g}K\left(\sin\frac{\alpha}{2}\right),$$

where K is the complete elliptic integral of the first kind, and α is the initial deflection.

(4) *Compute cases of interest to the client.* When the analytical solution cannot be used directly to give the desired assistance, and this is almost always the case, then we must compute a sufficient number of instances to enable us to interpret the solution. Before the development of high-speed computing machinery, computation was a long and tedious chore, something that was too often skimped.

Today there is a fast developing branch of mathematics that we might call computational mathematics, and a growing number of specialists in this field. Those who practice mathematics with this as their special field might be called *computer engineers*—certainly the term programmer is too narrow. This is a term suggested by Professor Tukey. In his words, "Computational mathematics recognizes one interpolating polynomial through a given set of points, and notes in passing that values of this polynomial can be found by [several different methods]. Computation engineering is concerned with the advantages and disadvantages of the various methods, and with the choice between them in particular situations" [8].

Of course the computer engineer must not stop there. His field of study must encompass the characteristics and special abilities of the various commercially available machines, the principles of the design of computing machines, and all developments in the use of computing machinery.

There is a tendency to use the high speed and immense size of these computers with little or no regard for efficiency. The computer engineer must combat this tendency. To him a computing machine is not a brain, but a supersonic moron. It must be told what to do every step of the way, but once it is told, it doesn't easily forget, and it is extremely obedient. The computer engineer must watch for spurious results or methods, and call attention of the client or a mathematical analyst to the difficulty. Very few investments lead to more embarrassment than computing voluminous data from the wrong set of formulas. We should see redder faces when an expensive computation is proved useless by a day or two of pencil and paper mathematics.

(5) *Interpret the results.* This activity is similar to problem recognition and understanding, and will almost always be performed by the same individual. In fact, often the groundwork for interpretation will be laid during the negotiation activity.

These five activities of the practicing mathematician are given grossly unequal emphasis in our training of mathematicians today. Most of our course work is geared toward Manipulation, with Computation coming more recently into the picture because of the vast number of computer engineers required today in business and industry. Some attempts at Formulation are included in our curricula, especially in applied mathematics courses but more often in nonmathematical courses. As for Negotiation and Interpretation, the almost com-

plete absence of anything resembling them from our mathematical curricula definitely handicaps newly trained mathematicians in their contacts with their clients.

Getting down to cases. A classical example of the practice of mathematics was the discovery of the planet Neptune by Leverrier [9]. The problem situation arose as observations of the erratic behavior of Uranus were recorded. Negotiation with astronomers led Leverrier to the hypothesis that another planet was causing the disturbance, and ruled out possible gravitational anomalies, ether resistance or disturbance by a comet. The mathematical model required was formulated and then manipulated until it was in a form suitable for computation. After the computations were completed, Leverrier interpreted the results to the German astronomer, Galle, as coordinates of a point in the sky where he would find the new planet. Galle found it within one degree of the predicted position.

Other examples of the practice of mathematics may be interesting. These examples are selected for their ease of explanation, not for their mathematical sophistication nor for their economic importance. No single example, or group of examples, could fairly picture the activity of the practicing mathematician.

Most of us have equated mathematics with what we refer to in this article as Manipulation, although many of us have now come to accept Computation as a legitimate mathematical occupation. The next illustration shows that Interpretation and Formulation may be all that is necessary in some situations.

A client was using a strange method for determining the buckling load of a stepped column, that is, one made up of several different uniform segments. It seems that there were two different ways to apply this method, but the client couldn't make the two answers jibe. Not only that, but neither answer was at all reasonable, and even averaging the two doubtful answers failed to improve them, strangely enough. First the mathematician wanted to know where the client found his method. He referred to an article in a reputable journal, and it turned out that as far as could be determined he was following the directions supplied in the article correctly. Careful study disclosed, however, that the article contained a set of three equations in only two unknowns, one of the alleged unknowns being a well-masked combination of the other two. Following this lead, a fundamental error was soon found hidden on the first page of the article. It was only necessary to point out the difficulty to the client, who was quite familiar with the manipulation required to get the corrected result that he needed.

A detective story writer could give some nice lilting title like "The Case of the Repressed Redundancy", or the "Case of the Curious Client", to this problem situation. It is quite fortunate that the redundant equations were incompatible, leading the curious client to suspect murder. Not all the errors in our voluminous printed literature lie so near the surface. But it is tempting for an engineer or other potential client to run down to the formula he wants and pay no

attention to where it came from. It seems that much valuable time is saved that way.

The next case is also a very simple one. The client complained that formulas for friction coefficients were not giving expected results in a certain case. He had a "dolly" for transporting aircraft sub-assemblies, which was equipped with V-grooved casters, and which was to be pulled along angle-iron "tracks". The client had to find the force necessary to pull the dolly along the track. At first he assumed rolling friction, and found a result which was absurdly low, but sliding friction gave a result far too high. After looking the problem over, the mathematician convinced himself and the client that rotation of the caster took place about some instantaneous axis which wasn't yet known, and whose determination was a part of the problem. Above this instantaneous axis the friction force opposed the forward motion of the caster, but below the instantaneous axis the friction force was actually in the direction of the motion. Points below the instantaneous axis move in the same fashion as points on the flange of a locomotive wheel—they actually move backward part of the time. Force and moment balance equations readily disclosed the force required and the position of the instantaneous axis.

Here again Manipulation and Computation played a minor role, but in this case conviction or Negotiation was an important part of the job. The client will not use the results obtained for him if he is not convinced of their validity. And he must be willingly convinced—not merely beaten with the force of authority. Both the mathematician and his client must agree that the problem as stated will lead to a conclusion in which some faith can be placed. And the more earth-shaking the conclusion, the more abiding the faith must be.

These problems are very simple ones, although they do look simpler after they are completed than they do before. There is a multitude of other problems —much more complex, more challenging. How should the throat of a supersonic wind tunnel be designed so that shock waves will not be generated by the walls of the tunnel itself? What temperatures are generated during machining? What is the best design for a shipping container? And there are many others.

How many clerks should be stationed behind the counter of a tool crib—taking into account the cost of clerks' idle time as well as the cost of mechanics' waiting time in the lines at the counter [10]? There are routines that can be used to solve linear programming problems on high-speed digital computing equipment. Can a much less expensive analog computer be used to solve such problems, and are there any other advantages in the use of analog equipment [11]? What change in performance of a given airplane can be expected if different engines are used?

To go on to even broader fields, consider that we human beings have had centuries of experience in detecting signals in the presence of noise. We can see a tomato worm on a tomato plant in spite of his protective coloring, we may hear a trout splashing in a stream even above the noise of the stream, we hear radio signals even when there is static. How can this human experience be built into

machinery, so that signals can be detected mechanically, electronically, mathematically, automatically—in the presence of noise? Progress continues in this important field [12].

Another very important problem facing us today is that of educating a computing machine. We must take these arithmetic strong-backs and teach them to communicate more intelligently with their masters, so that hordes of people aren't required to boss them. We have already taught them algebra [13], but this is just a beginning.

Diversity in the training of mathematicians. For the good of our technology, it is imperative that the position of mathematicians be established as basic. They should be used continuously and regularly, and not just to provide occasional flashes of genius when all else fails. Until this happy day dawns, the position of the practicing mathematician will be somewhat less than clear, and much of his best work—that in which his ideas and methods provoke an evolution in the designs and decisions of his clientèle—will be unrecognized and unrewarded.

We can speed the coming of that day if we develop more mathematicians whose thought is problem-centered rather than method-centered. We need mathematicians whose training is broad enough so that they can give sound advice in a wide variety of problem situations. Furthermore, we should encourage more diversity in the curricula of our mathematicians. I am in good company in making this suggestion [8], and the same conclusion, by reasoning from a different direction, has been much more forcefully drawn by Oakley [14].

As it is now, many of those who practice mathematics today are converted from other fields, such as physics, electrical and mechanical engineering, and chemistry. Perhaps conversions of this kind provide the most practicable way to get broadly trained mathematicians, since many courses in engineering and science are accessible only through long chains of prerequisites, some of them realistic and some not so realistic. On the other hand, we would do well to consider the needs of the practicing mathematician when we find a student who wants to build up a stronger-than-usual background in a cognate field.

References

1. T. C. Fry, Industrial mathematics, this MONTHLY, vol. 48, no. 6, part II, 1941, p. 5.
2. Th. von Karman, Aerodynamics, Cornell University, 1954, p. 2.
3. George F. Nordenholt, A graduate can measure a bottle, Product Engineering, April 1953, p. 135.
4. Anonymous, Business Week, June 30, 1951, p. 45.
5. J. D. Williams, The Compleat Strategyst, New York, 1954, p. 215.
6. A. N. Whitehead, Science in the Modern World, New York, 1931, p. 32.
7. Rufus Oldenburger, Mathematical Engineering Analysis, New York, 1950.
8. J. W. Tukey, Mathematical consultants, computational mathematics, and mathematical engineering, this MONTHLY, vol. 62, 1955, p. 566.
9. Louis Figuier, Principales Découvertes Scientifiques Modernes, vol. IV, Paris, 1858, pp. 291–319. This contemporary account of the discovery is the one which we follow.
10. Georges Brigham, On a congestion problem in an aircraft factory, J. Operations Res. Soc. Amer., vol. 3, 1955, p. 412.

11. C. M. Ablow and Georges Brigham, An analog solution of programming problems, J. Operations Res. Soc. Amer., vol. 3, 1955, p. 388.

12. W. M. Stone, On the statistical theory of detection of a randomly modulated carrier, J. Appl. Phys., vol. 24, 1953, p. 935.

13. Mandalay Grems and R. E. Porter, A truly automatic computing system, Proc. Western Joint Computer Conference, Feb. 1956.

14. C. O. Oakley, Engineering for mathematicians, J. Engineering Education, vol. 47, 1956, p. 292.

ON HIGHER ORDERED DIFFERENTIATION

DAVE PANDRES, JR., University of Texas

1. Introduction. Courant in [1] gives Leibniz' rule for finding the Nth derivative of the product of two functions, and then remarks that no such easily remembered law has been found for the repeated differentiation of the compound function $Y=F[U(X)]$. Rather complicated formulae for the Nth derivative of a compound function have been found by Schlömilch [2], Faa' di Bruno [3], and McKiernan [4]. In this paper, however, a simple and easily remembered law for repeated differentiation is given which is valid not only for a compound function, but also for the more general composite function $Y=F[U_1(X), \cdots, \cdots, U_Q(X)]$. Partially to illustrate the power of this law, but more particularly because the problem is of some interest in its own right, a formula is derived for writing the power series representation of a meromorphic function of finite genre in terms of its Weierstrass product representation.

In each of the results given in this paper, a certain matrix of simple structure plays a prominent role. It is shown that the characteristic polynomial of this matrix has interesting properties, especially with regard to differentiation processes.

It may be worthy of note that in the formula for the repeated differentiation of a composite function, the elements of the matrix are operational quantities similar in form to the familiar chain rule for first ordered differentiation. Thus, the formula given here may be regarded as a chain rule for higher ordered differentiation.

2. The matrix Δ_N. The symbol Δ_N denotes the Nth ordered matrix

$$\begin{bmatrix} D_1 & -1 & 0 & 0 & 0 & \cdots & 0 \\ D_2 & D_1 & -2 & 0 & 0 & \cdots & 0 \\ D_3 & D_2 & D_1 & -3 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & & 1-N \\ D_N & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \begin{matrix} \\ \\ \\ \\ \\ \\ \\ \\ D_3 \quad D_2 \quad D_1 \end{matrix} .$$

The elements $D_1, \cdots, D_N$ will be assigned meanings appropriate to the particular topics under discussion. In some cases these elements will be operational in nature, while in others they will be purely algebraic. The symbol $|\Delta_N|$ denotes the determinant of Δ_N. The zero'th ordered determinant $|\Delta_0|$ is defined as unity.

3. The *N*th derivative of a composite function. Let the function $Y = F(U_1, \cdots, U_Q)$ be continuous together with its mixed partial derivatives of the first N orders with respect to $U_1, U_2, \cdots, U_Q$. Let $U_1(X), \cdots, U_Q(X)$ be continuous together with their first N derivatives with respect to X. It is then possible to state the following theorem.

THEOREM 1. *The Nth derivative of Y with respect to X is given by*

$$\frac{d^N Y}{dX^N} = |\Delta_N| F(U_1, \cdots, U_Q),$$

where the elements $D_1, \cdots, D_N$ *of* Δ_N *are defined by the relation*

$$D_i = \frac{1}{(i-1)!} \sum_{h=1}^{Q} \frac{d^i U_h}{dX^i} \frac{\partial}{\partial U_h} .$$

Proof. For $N=1$, the theorem is merely an operational statement of the familiar chain rule for first ordered differentiation. In order to complete a proof by induction, it will now be assumed that the theorem is true for $N=1, \cdots, K$, and it will be shown that it must then be true also for $N=K+1$. Upon expanding $|\Delta_K|$ by cofactors of its last row, one obtains the recurrence relation

$$|\Delta_K| = \sum_{i=1}^{K} \frac{(K-1)!}{(K-i)!} D_i |\Delta_{K-i}| .$$

Then, substituting the operational definitions for the elements D_i gives

$$|\Delta_K| = \sum_{i=1}^{K} \sum_{h=1}^{Q} \frac{(K-1)!}{(K-i)!(i-1)!} \frac{d^i U_h}{dX^i} \frac{\partial}{\partial U_h} |\Delta_{K-i}| .$$

If both sides of this relation are allowed to operate upon Y, the result is

$$\frac{d^K Y}{dX^K} = \sum_{i=1}^{K} \sum_{h=1}^{Q} \frac{(K-1)!}{(K-i)!(i-1)!} \frac{d^i U_h}{dX^i} \frac{\partial}{\partial U_h} \frac{d^{K-i} Y}{dX^{K-i}}.$$

Then, differentiation with respect to X yields the expression

$$\frac{d^{K+1} Y}{dX^{K+1}} = \sum_{i=1}^{K} \sum_{h=1}^{Q} \frac{(K-1)!}{(K-i)!(i-1)!} \left[\frac{d^{i+1} U_h}{dX^{i+1}} \frac{\partial}{\partial U_h} \frac{d^{K-i} Y}{dX^{K-i}} + \frac{d^i U_h}{dX^i} \frac{\partial}{\partial U_h} \frac{d^{K-i+1} Y}{dX^{K-i+1}} \right].$$

This result may be written in the alternative form

$$\frac{d^{K+1} Y}{dX^{K+1}} = \sum_{i=1}^{K} \frac{(K-1)!}{(K-i)!} (iD_{i+1} | \Delta_{K-i}| + D_i | \Delta_{K-i+1}|) Y.$$

Then, upon collecting like terms, one obtains

$$\frac{d^{K+1} Y}{dX^{K+1}} = \sum_{i=1}^{K+1} \frac{K!}{(K+1-i)!} D_i | \Delta_{K+1-i}| Y.$$

The expression operating upon Y is precisely $|\Delta_{K+1}|$ expanded by cofactors of its last row. Therefore, $d^{k+1} Y/dX^{k+1} = |\Delta_{K+1}| Y$, and hence, by the postulate of mathematical induction, $d^N Y/dX^N = |\Delta_N| Y$ for all positive integral values of N.

COROLLARY. *Let $Y = \epsilon^{U(X)}$ where $U(X)$ is continuous together with its first N derivatives. Then*

$$\frac{d^N Y}{dX^N} = | \Delta_N | \epsilon^{U(X)},$$

where the elements $D_1, \cdots, D_N$ of Δ_N are defined by the relation

$$D_i = \frac{1}{(i-1)!} \frac{d^i U}{dX^i}.$$

Proof. The corollary follows directly from Theorem 1. Since each of the derivatives of ϵ^U with respect to U is precisely ϵ^U itself, one may expand $|\Delta_N|$, allow the expansion to operate upon ϵ^U, and then place the result again in the form of a determinant by merely omitting the partial differentiation symbols in the relation defining $D_1, \cdots, D_N$.

4. Meromorphic functions of finite genre. A function which is analytic, except for a finite number of poles within any finite region, at every finite point of the complex plane is said to be a meromorphic function. Let $F(Z)$ be a meromorphic function whose zeros (excluding any which may lie at the origin) are at $a_1, a_2, \cdots$. Let the zero at a_i be of order m_i. Poles are regarded as zeros of negative order. Let the zero at the origin be of order m_0. It then follows by the Weierstrass factor theorem [5]. that $F(Z)$ may be represented at every point for which it is defined by the following infinite product:

$$F(Z) = \epsilon^{G(Z)} Z^{m_0} \prod_{i=1}^{\infty} \left[\left(1 - \frac{Z}{a_i}\right) \epsilon^{H_i(Z)} \right]^{m_i},$$

where

$$H_i(Z) = \sum_{n=1}^{k_i} \frac{1}{n} \left(\frac{Z}{a_i}\right)^n,$$

$G(Z)$ is analytic throughout the entire finite plane, and the k_i are chosen so as to ensure the convergence of the product.

It is shown in [6] that if K is the smallest nonnegative integer such that the summation $\sum_{i=1}^{\infty} m_i/a_i^{K+1}$ is absolutely convergent, then it is sufficient to choose $k_i = K$ for all values of i. If with this choice for k_i the function $G(Z)$ is of (finite) degree K', then the function $F(Z)$ is said to be of finite genre, the genre P being the larger of the two nonnegative integers K and K'. In other words the genre of the function $F(Z)$ is the highest (necessary) power of Z in any exponential function within or without the product sign. If $F(Z)$ is of genre P, therefore, it may be represented by the Weierstrass product given above, except that the functions $G(Z)$ and $H_i(Z)$ will now be polynomials of (formal) degree P. That is,

$$G(Z) = \sum_{n=0}^{P} C_n Z^n, \quad H_i(Z) = \sum_{n=1}^{P} \frac{1}{n} \left(\frac{Z}{a_i}\right)^n.$$

5. The series expansion of a function of finite genre. Let $F(Z)$ be the function of finite genre described above. Let its representation by a power series be

$$F(Z) = Z^{m_0} \sum_{n=0}^{\infty} A_n Z^n.$$

It is then possible to state the following theorem.

THEOREM 2.

$$A_N = \frac{1}{N!} \left| \Delta_N \right| \epsilon^{C_0},$$

where the elements $D_1, \cdots, D_N$ *of* Δ_N *are defined by*

$$D_K = KC_K, \quad K \leqq P; \qquad D_K = -\sum_{i=1}^{\infty} \frac{m_i}{a_i^K}, \qquad K > P.$$

Proof. Since $F(Z) = \epsilon^{\log F(Z)}$, one may write at once from the corollary to Theorem 1,

$$\frac{d^N F(Z)}{dZ^N} = \left| \Delta_N \right| \epsilon^{\log F(Z)}, \quad \text{where} \quad D_K = \frac{1}{(K-1)!} \frac{d^K \log F(Z)}{dZ^K}.$$

But the logarithm of the Weierstrass product representation of $F(Z)$ is given by

$$\log F(Z) = \sum_{n=0}^{P} C_n Z^n + \sum_{i=1}^{\infty} m_i \left[\log\left(1 - \frac{Z}{a_i}\right) + \sum_{n=1}^{P} \frac{1}{n}\left(\frac{Z}{a_i}\right)^n \right].$$

(Here, it is temporarily assumed that $m_0 = 0$.) Upon differentiating K times with respect to Z, one obtains

$$\frac{d^K \log F(Z)}{dZ^K} = \sum_{n=K}^{P} \frac{n!}{(n-K)!} C_n Z^{n-K} + \sum_{i=1}^{\infty} m_i \left[-\frac{(K-1)!}{(a_i - Z)^K} + \sum_{n=K}^{P} \frac{(n-1)!}{(n-K)!} \frac{Z^{n-K}}{a_i^n} \right].$$

Evaluating this expression for $Z = 0$, and dividing by $(K-1)!$ gives the result

$$\frac{1}{(K-1)!} \frac{d^K \log F(Z)}{dZ^K}\bigg|_{Z=0} = \begin{cases} -\sum_{i=1}^{\infty} \frac{m_i}{a_i^K}, & K > P \\ KC_K, & K \leqq P. \end{cases}$$

Now, $\epsilon^{\log F(Z)}$ evaluated for $Z = 0$ is simply ϵ^{C_0}, so

$$\frac{d^N F(Z)}{dZ^N}\bigg|_{Z=0} = |\Delta_N| \epsilon^{C_0}.$$

It follows at once from Taylor's theorem that the coefficient of the Nth term in the series expansion of $F(Z)$ is given by

$$A_N = \frac{1}{N!} |\Delta_N| \epsilon^{C_0},$$

where

$$D_K = -\sum_{i=1}^{\infty} \frac{m_i}{a_i^K}, \quad K > P; \qquad D_K = KC_K, \quad K \leqq P.$$

Now it is obvious that if m_0 is not zero, then each term in the series expansion is merely multiplied by Z^{m_0}. Thus, the coefficient of the Nth term is unchanged, although the power of the term will now be $N + m_0$.

It should be noted that in the proof given here, it is not necessary to assume that the multiplicities m_i and m_0 are integers, or even that they are real, although these assumptions are satisfied if the function $F(Z)$ is meromorphic. Thus, the theorem given here is valid not only for meromorphic functions, but also for any other functions which may be expressed in the *form* of a Weierstrass product for a meromorphic function of finite genre. If, however, multiplicities are present which are not real integers, then it must be recognized that the resulting series will represent the principal branch of the function.

6. The characteristic polynomial of the matrix Δ_N. Let the elements of Δ_N be complex numbers, and let $P_N(X)$ denote the characteristic polynomial of Δ_N. It is then possible to state the following theorem.

THEOREM 3.

$$\frac{dP_N(X)}{dX} = NP_{N-1}(X).$$

Proof. The characteristic polynomials associated with Δ_1 and Δ_2 are

$$P_1(X) = X - D_1,$$

$$P_2(X) = \begin{vmatrix} X - D_1 & 1 \\ -D_2 & X - D_1 \end{vmatrix} = X^2 - 2D_1X + D_1^2 + D_2.$$

The derivative of $P_2(X)$ is $2P_1(X)$. Thus, the theorem is seen to be true for $N=2$. It is also formally true for $N=1$, since $P_0(X)$ is unity by definition. In order to complete a proof by induction, it will now be shown that if the theorem is true for $N=1, \cdots, K-1$, then it is also true for $N=K$.

In order to obtain the characteristic polynomial associated with the matrix Δ_K, it is only necessary to substitute for D_i the quantity $(\delta_i^1 X - D_i)$, and for the elements $-1, -2, \cdots, 1-K$, their absolute values, and to expand the result as a determinant. Here, the symbol δ_i^1 denotes the so-called Kronecker delta, which by definition assumes the value unity for $i=1$ and the value zero for $i \neq 1$. Thus, expanding the determinant by cofactors of its last row gives

$$P_K(X) = \sum_{i=1}^{K} \frac{(K-1)!}{(K-i)!}(-1)^{i-1}(\delta_i^1 X - D_i)P_{K-i}(X).$$

Then, upon differentiating with respect to X, one obtains

$$P'(X) = \sum_{i=1}^{K} \frac{(K-1)!}{(K-i)!}(-1)^{i-1}[\delta_i^1 P_{K-i}(X) + (\delta_i^1 X - D_i)P'_{K-i}(X)].$$

But, by assumption $P'_N(X) = NP_{N-1}(X)$, for $N=1, 2, \cdots, K-1$, so

$$P'_K(X) = \sum_{i=1}^{K-1} \frac{(K-1)!}{(K-i)!}(-1)^{i-1}[\delta_i^1 P_{K-i}(X) + (\delta_i^1 X - D_i)(K-i)P_{K-i-1}(X)].$$

Then, because of the definition of the Kronecker delta,

$$P_K(X) = P_{K-1}(X) + (K-1)\sum_{i=1}^{K-1} \frac{(K-2)!}{(K-1-i)!}(-1)^{i-1}(\delta_i^1 X - D_i)P_{K-i-1}(X)$$

But the summation of the right is precisely the characteristic polynomial of Δ_{K-1}. Thus, $P'_K(X) = KP_{K-1}(X)$. And, by the postulate of mathematical induction, $P'_N(X) = NP_{N-1}(X)$ for all N.

COROLLARY. *Let the symbol C_K denote the coefficient of X^K in $P_N(X)$. It then follows that*

$$C_K = (-1)^{N-K}\frac{N!}{(N-K)!K!}\,|\,\Delta_{N-K}\,|\,.$$

Proof. From Theorem 3, one may write at once for $K \leqq N$,

$$\frac{d^K P_N(X)}{dX^K} = \frac{N!}{(N-K)!} P_{N-K}(X).$$

But, by definition

$$P_N(X) = \sum_{n=0}^{N} C_n X^n.$$

So it follows that

$$\frac{d^K P_N(X)}{dX^K} = \sum_{n=K}^{N} \frac{n!}{(n-K)!} C_n X^{n-K}.$$

Equating the two expressions for $d^K P_N(X)/dX^K$ and setting $X=0$ gives

$$K!C_K = \frac{N!}{(N-K)!} P_{N-K}(0).$$

But, $P_{N-K}(0)$ is simply the constant term in $P_{N-K}(X)$, so its value is $(-1)^{N-K}$ times $|\Delta_{N-K}|$. Thus, it follows that

$$K!C_K = \frac{N!}{(N-K)!} (-1)^{N-K} |\Delta_{N-K}|,$$

and finally that

$$C_K = (-1)^{N-K} \frac{N!}{(N-K)!K!} |\Delta_{N-K}|.$$

References

1. R. Courant, Differential and Integral Calculus, 2nd ed., New York, 1937, vol. I, p. 202.
2. O. Schlömilch, Compendium der Höheren Analysis, Brunswick, 1879, vol. 2, p. 4.
3. F. Faa' di Bruno, Annales de Tortolini, 1855 and Théorie des Formes Binaires.
4. M. McKiernan, On the *N*th derivative of composite functions, this MONTHLY, vol. 63, 1956, pp. 331–333.
5. E. T. Whittaker and G. N. Watson, A Course of Modern Analysis, Cambridge, 1935, p. 137.
6. E. Borel, Leçons sur les Fonctions Entières, Paris, 1900, p. 24.

ON CERTAIN MATRIX EQUATIONS

T. J. PIGNANI, University of North Carolina

This paper exhibits some properties of matrix solutions for certain matrix equations, where the elements in the coefficient matrices range over the field of complex numbers. In particular, the first section gives the number of parameters along with a minimum and maximum rank of the matrix solution for $AYB=C$. The second section gives the minimum number of parameters in the solution for matrix equations which are of a more general type, and the third section is devoted to remarks.

Throughout, capital letters denote matrices, while lower case letters denote scalars. The symbol A^I means the inverse of the matrix A, when A is nonsingular, and without confusion, the symbol 0 is used to designate the matrix with every element zero or the scalar of value zero as the case may be. The collection of symbols $A(m, n, a)$ indicates that the matrix A has m rows, n columns and is of rank a.

1. Concerning the matrix equation $AYB=C$. The following theorem establishes the minimum rank of the matrix solution Y of

$$AYB = C, \tag{1}$$

where $A(m, n, a)$, $0<a\leqq\min(m, n)$, $B(k, t, b)$, $0<b\leqq\min(k, t)$, $C(m, t, c)$, $0\leqq c\leqq\min(m, t)$, while Y is an unknown n by k matrix.

THEOREM 1. *If $c\leqq a$, $c\leqq b$, then the rank of the matrix solution Y of (1) is not less than c, providing necessary and sufficient conditions [1] are satisfied for the existence of a solution.*

Proof. The rank of the product AYB does not exceed the rank of its factors [2, p. 42], but this rank is c. Hence the rank of Y is not less than c.

A well known result [2, p. 11] is that if A is a matrix as described above, there exist nonsingular matrices L_A, R_A such that $L_AAR_A=E_a$, where E_a is a matrix whose elements e_{ij} have the following values: $e_{ij}=1$, $(i=j=1, \cdots, a)$, $e_{ij}=0$, $i\neq j$, $(i=1, \cdots, m)$, $(j=1, \cdots, n)$.

THEOREM 2. *If $c\leqq a$, $c\leqq b$ and the last $(m-a)$ rows and $(t-b)$ columns of L_ACR_B consist of zero elements, then there exists an $(nk-ab)$-parameter family of solutions for Y and of these, there is at least a p^*-parameter family of solutions each of maximum rank r, where p^* and r are given below.*

Proof. Equation (1) is rewritten as

$$E_aZE_b = D, \tag{2}$$

where $Z=R_A^IYL_B^I$, $D=L_ACR_B$. It is observed that with the hypotheses of this theorem, equation (2) satisfies necessary and sufficient conditions [1] for a matrix solution to exist, hence a matrix solution Y of (1) exists. The product

E_aZE_b is an m by t matrix of the form

$$\left(\begin{array}{ccc|c} z_{11} & \cdots & z_{1b} & \\ \cdots & \cdots & \cdots & 0 \\ z_{a1} & \cdots & z_{ab} & \\ \hline & 0 & & 0 \end{array}\right),$$

and by hypothesis

$$D = L_ACR_B = \left(\begin{array}{ccc|c} d_{11} & \cdots & d_{1b} & \\ \cdots & \cdots & \cdots & 0 \\ d_{a1} & \cdots & d_{ab} & \\ \hline & 0 & & 0 \end{array}\right).$$

Hence ab of the elements in Z are uniquely determined while $(nk-ab)$ of its elements are parameters. Further the matrix solution Y of (1) contains these parameters, since $Y=R_AZL_B$.

To show that, of these solutions, there is at least a p^*-parameter family of solutions each of maximum rank r, first transform Z, by use of elementary transformations, to an equivalent matrix X—in rank—so that the nonvanishing determinant of order c appears in the first c rows and c columns of X. Assign the following nonzero values to the parametric elements x_{pq} of X in each of the following cases:

I. If $n \leqq k$

(A) and $n=a$, either

1. $a>c$ and $k>b$, let $x_{c+i,b+i}=\lambda_i$, $(i=1, \cdots, p^*)$, where $p^*=\min(a-c, k-b)$, otherwise $x_{pq}=0$. In this case $r=c+p^*$.
2. or $a=c$, $k>b$, let $x_{pq}=0$. In this case $p^*=0$, $r=n$. Note this case does not exist when $n=k$.
3. or $a \geqq c$, $k=b$. In this case there are no parameters, hence $p^*=0$, $r=c$.

(B) or $n>a$, and either

1. $c+n-a \geqq b$, and
   (a) $a>c$, let $x_{a+i,c+i}=\lambda_i$, $(i=1, \cdots, n-a)$, $x_{c+j,c+n-a+j}=\lambda_j$, $(j=1, \cdots, a-c)$, otherwise $x_{pq}=0$. In this case $p^*=n-c$, $r=n$.
   (b) or $a=c$, let $x_{a+i,c+i}=\lambda_i$, $(i=1, \cdots, n-a)$, otherwise $x_{pq}=0$. In this case $p^*=n-a$ and $r=n$.
2. or $c+n-a<b$, and
   (a) $a>c$ and $k>b$, let $x_{a+i,c+i}=\lambda_i$, $(i=1, \cdots, n-a)$, $x_{b+j,c+j}=\lambda_j$, $(j=1, \cdots, s)$, where $s=\min(a-c, k-b)$, otherwise $x_{pq}=0$. In

this case $p^*=n-a+s$ and $r=c+p^*$.

(b) or $a=c$ and $k\geqq b$ or $a\geqq c$ and $k=b$, let $x_{a+i,c+i}=\lambda_i$, $(i=1, \cdots, n-a)$, otherwise $x_{pq}=0$. In this case $p^*=n-a$ and $r=n$ or $r=c+p^*$ respectively.

II. Or $n\geqq k$. This is treated as Case I by considering the transpose of (1), namely, $B^TY^TA^T=C^T$, and various cases follow in a manner entirely similar to that above. Since R_A and L_B are nonsingular matrices and $Y=R_AZL_B$, the matrix Y is of the same rank as Z.

It is worth noting that the hypothesis on the matrix L_ACR_B is a necessary condition for the systems to be consistent.

Certain cases of these results are essential in the study of differential systems with interface conditions as indicated in [3].

2. Certain general matrix equations. Consider the matrix equation

$$\sum_{h=1}^{\alpha} A_hY_hB_h = C, \tag{3}$$

where $A_h(m, n_h, a_h)$, $0<a_h\leqq\min(m, n_h)$, $B_h(k_h, t, b_h)$, $0<b_h\leqq\min(k_h, t)$, $C(m, t, c)$, $0\leqq c\leqq\min(m, t)$, while Y_h are unknown n_h by k_h matrices. Denote with "a" and "b" the ranks of the matrices such that $a_hb_h\leqq ab$, $(h=1, \cdots, \alpha)$. Also "condition LR" means $L^I_{A_1}=L^I_{A_2}=\cdots=L^I_{A_\alpha}$, $R^I_{B_1}=R^I_{B_2}=\cdots=R^I_{B_\alpha}$.

THEOREM 3. *If in* (3) *condition LR holds, $C\neq 0$, the last $(m-a)$ rows and $(t-b)$ columns of $L_{A_h}CR_{B_h}$ consist of zero elements, there is at least an $(n_hk_h-a_hb_h)$-parameter family of solutions for each Y_h.*

Proof. Equation (3) is rewritten as

$$\sum_{h=1}^{\alpha} E_{a_h}Z_hE_{b_h} = D, \tag{4}$$

where $Z_h=R^I_{A_h}Y_hL^I_{B_h}$, $D=L_{A_h}CR_{B_h}$. With close observation, it is seen equation (4) satisfies necessary and sufficient conditions [1] for solutions, Z_h, to exist under the hypotheses of this theorem. In each matrix Z_h, at least $(n_hk_h-a_hb_h)$ of the elements cannot be determined. Since $Y_h=R_{A_h}Z_hL_{B_h}$, then Y_h contains these parameters.

THEOREM 4. *If in* (3) *condition LR holds, $C=0$, $ab<\sum_{h=1}^{\alpha}n_hk_h$, there is at least an $(n_hk_h-a_hb_h)$-parameter family of solutions for Y_h not all zero.*

In case $Y=Y_h$, $n=n_h$, $k=k_h$, $(h=1, \cdots, \alpha)$, then (3) becomes

$$\sum_{h=1}^{\alpha} A_hYB_h = C. \tag{5}$$

Under the same hypotheses as in Theorem 2 and in a manner entirely similar to that above follows

THEOREM 5. *There is at least an* $(nk-ab)$*-parameter family of solutions.*

THEOREM 6. *If in* (5), *condition LR holds,* $C=0$, $nk \neq ab$, *there is at least an* $(nk-ab)$*-parameter family of nonzero solutions.*

3. Remarks. In a recent and interesting note [4], the existence and number of solutions were determined for the matrix equation $AX=B$, where the elements in A and B belong to a finite field. The number of solutions is given in terms of the order of the field and the number of parameters in the solution. Further, the author gives the number of solutions for the matrix equation $UA+BV=C$, if solutions exist.

In my work presented above and following Hodges [4], the number of solutions of (1), (3), and (5) is unbounded. If one restricts the elements of the matrices in (1) to range over a finite field of order f, then the number of solutions is f^{nk-ab}.

References

1. W. E. Roth, Direct product matrices, Bull. Amer. Math. Soc., vol. 40, 1934, pp. 461–468.
2. C. C. MacDuffee, The Theory of Matrices, New York, 1946.
3. T. J. Pignani and W. M. Whyburn, Differential systems with interface and general boundary conditions, J. Elisha Mitchell Sci. Soc., vol. 72, 1956, pp. 1–14.
4. J. H. Hodges, The matrix equation $AX=B$ in a finite field, this MONTHLY, vol. 63, 1956, pp. 243–244.

THE DIRICHLET SERIES TRANSFORMATION*

J. RAY HANNA, University of Wichita

1. Introduction. In references [3] and [4], Tomlinson Fort introduced the Dirichlet series transformation as a tool for solving difference equations. We define the transformation by the relationship,

$$\mathfrak{D}\{F(t)\} = f(s) = \sum_{t=0}^{\infty} a^{-st}F(t), \qquad a > 1. \tag{1}$$

$F(t)$ is assumed to be a function of a discrete variable t, where t takes on nonnegative integral values. Deviating from Fort's definition, s is assumed complex.

If $F(t)=O(k^t)$ for $t>t_0$, then (1) is absolutely convergent when $a^{\mathrm{Re}(s)}>k$, (k, a positive constant).

2. Inversion. Inversion of (1) can be accomplished by differentiation or contour integration if s is a complex variable. It can be shown that $F(t)$ is given by the theorem.

* Excerpt from a thesis written under the direction of Professor J. R. Britton, University of Colorado.

INVERSION THEOREM. *Assume that* $f(s)=\sum_{t=0}^{\infty} a^{-st}F(t)$, $a>1$. (a) *If* $f(s)$ *is such that* $f\{-(\operatorname{Ln} w)/(\operatorname{Ln} a)\}$ *is single-valued and analytic inside some circle* C *with center* $w=0$, *then*

$$F(t)=\lim_{s\to\infty}\frac{1}{t!}\frac{d^t}{dw^t}\,f\left(\frac{-\operatorname{Ln} w}{\operatorname{Ln} a}\right)\Bigg]_{w=a^{-s}}$$

(b) *If* $f(s)$ *is such that* $f\{(\operatorname{Ln} z)/(\operatorname{Ln} a)\}$ *is single-valued and analytic outside some circle* C *with center* $z=0$, *then*

$$F(t)=\frac{1}{2\pi i}\int_C z^{t-1}f\{(\operatorname{Ln} z)/(\operatorname{Ln} a)\}\,dz=\sum R_n,$$

where R_n *are the residues at the singularities,* Z_n, *of the integrand inside* C. (*We use* Ln *to stand for the principal value of the logarithm.*)

In addition to the inversion theorem, the use of a table of transforms relating $f(s)$ and $F(t)$ is desirable for the problem of inversion. If this latter procedure is used, it is imperative that for a given $f(s)$ there be associated a function $F(t)$ which is unique. Uniqueness can be justified, if $F(t)$ is of exponential order, by a process described in [2, pp. 166–167].

3. Derivatives of transforms. If $F(t)=O(k_0^t)$ and $a^{\operatorname{Re}(s)}\geqq k$ where $k>k_0$, then $|a^{-st}|\,|F(t)|<K(k_0k^{-1})^t$. K and k_0 are positive constants. The function $(k_0k^{-1})^t$ is independent of s and its sum converges over the interval $(0,\infty)$. Therefore, (1) is uniformly convergent with respect to s. The Cauchy-Riemann equations are satisfied with $a^{-st}F(t)$. Hence termwise differentiation of (1) is permissible. Therefore,

$$(d/ds)f(s)=\sum_{t=0}^{\infty}a^{-st}[-tF(t)]\operatorname{Ln} a.$$

If $t^{n-1}F(t)$ satisfies the conditions imposed on $F(t)$, then

$$(d^n/ds^n)f(s)=\sum_{t=0}^{\infty}a^{-st}[(-t)^nF(t)](\operatorname{Ln} a)^n,$$

or

$$\mathfrak{D}\{t^nF(t)\}=(-1)^n/(\operatorname{Ln} a)^n\cdot(d^n/ds^n)f(s).$$

4. Translation of the transform $f(s)$. Let $F(t)$ in (1) be of exponential order.

$$f(s+\sigma)=\sum_{t=0}^{\infty}a^{-(s+\sigma)t}F(t),\ \text{where } \sigma \text{ is complex.}$$

$$f(s+\sigma)=\sum_{t=0}^{\infty}a^{-st}a^{-\sigma t}F(t)=\mathfrak{D}\{a^{-\sigma t}F(t)\},\ \text{when } a^{\operatorname{Re}(s)+\operatorname{Re}(\sigma)}>k.$$

5. Convolution. Assume $F(t)$ and $G(t)$ are two functions, each of exponential order. Then $\mathfrak{D}\{F(t)\}$ and $\mathfrak{D}\{G(t)\}$ converge absolutely. The Cauchy product of the two series,

$$f(s)g(s) = \sum_{t=0}^{\infty} a^{-st}F(t) \sum_{t=0}^{\infty} a^{-st}G(t),$$

may be expressed as

$$f(s)g(s) = \sum_{t=0}^{\infty} a^{-st}\left[\sum_{\tau=0}^{t} F(\tau)G(t-\tau)\right],$$

or

$$f(s)g(s) = \sum_{t=0}^{\infty} a^{-st}\left[\sum_{\tau=0}^{t} G(\tau)F(t-\tau)\right].$$

Therefore, we may express

$$f(s)g(s) = \mathfrak{D}\{F(t) * G(t)\} = \mathfrak{D}\{G(t) * F(t)\}, \tag{2}$$

where the combination of $F(t)$ and $G(t)$ appearing inside the sum is called the convolution of these functions. It is understood that

$$F(t) * G(t)] = \sum_{\tau=0}^{t} F(\tau)G(t-\tau).$$

With $G(t)=1$, in (2), it follows that

$$\frac{a^s}{a^s-1}f(s) = \mathfrak{D}\{1 * F(t)\} = \mathfrak{D}\left\{\sum_{\tau=0}^{t} F(\tau)\right\}. \tag{3}$$

If the process of (3) is continued n times,

$$\frac{a^{ns}}{(a^s-1)^n}f(s) = \mathfrak{D}\left\{\sum_{\tau=0}^{t}\sum_{\sigma=0}^{\tau}\cdots\sum_{\alpha=0}^{\beta} F(\alpha)\right\}, \tag{4}$$

where n indicates the number of summations.

6. Applications. The transform method, used by Fort [3] for the solution of a difference equation, can be applied to systems of linear difference equations with constant coefficients. This method is useful for the solution of certain summation equations also. Consider the equation,

$$\sum_{\tau=0}^{t} \tau^{(k)}F(t-\tau) = (t+n)^{(m)}, \qquad 1 \leqq n \leqq m-k. \tag{5}$$

(The notation $t^{(q)}$ represents the factorial polynomial $t(t-1)\cdots(t-q+1)$.) The transformed equation of (5) is

$$\frac{k!a^s}{(a^s-1)^{k+1}}f(s)=\frac{m!a^{(n+1)s}}{(a^s-1)^{m+1}},$$

or

$$f(s)=\frac{m!a^{ns}}{k!(a^s-1)^{m-k}}.$$

Since it can be shown that

$$\mathfrak{D}\{(t+p)^{(q)}\}=\frac{q!a^{(p+1)s}}{(a^s-1)^{q+1}},\text{ if } p\leqq q,$$

the solution of the equation follows immediately.

$$F(t)=\frac{m!}{k!(m-k-1)!}(t+n-1)^{(m-k-1)}.$$

The summation of finite series may be accomplished by employing results suggested by the convolution theorem. We observe that (3) and (4) may be written in the form,

$$\sum_{\tau=0}^{t}F(\tau)=\mathfrak{D}^{-1}\left\{\frac{a^s}{a^s-1}f(s)\right\},\tag{6}$$

and

$$\sum_{\tau=0}^{t}\sum_{\sigma=0}^{\tau}\cdots\sum_{\alpha=0}^{\beta}F(\alpha)=\mathfrak{D}^{-1}\left\{\frac{a^{ns}}{(a^s-1)^n}f(s)\right\}.\tag{7}$$

($\mathfrak{D}^{-1}\{f(s)\}$ is the inversion of $f(s)$ and is equal to $F(t)$.)

The cosine series, $1+\cos g+\cos 2g+\cdots+\cos tg$, may be summed by formula (6).

$$\sum_{\tau=0}^{t}\cos g\tau=\mathfrak{D}^{-1}\left\{\frac{a^s}{a^s-1}\,\frac{a^s}{2}\left[\frac{1}{a^s-e^{ig}}+\frac{1}{a^s-e^{-ig}}\right]\right\}$$

$$=\mathfrak{D}^{-1}\left\{\frac{a^{2s}}{2(a^s-1)(a^s-e^{ig})}+\frac{a^{2s}}{2(a^s-1)(a^s-e^{-ig})}\right\}.$$

Since it can be shown that

$$\mathfrak{D}\left\{\frac{b^{t+1}-1}{b-1}\right\}=\frac{a^{2s}}{(a^s-1)(a^s-b)},$$

then

$$\mathfrak{D}^{-1}\left\{\frac{a^{2s}}{2(a^s-1)(a^s-e^{ig})}+\frac{a^{2s}}{2(a^s-1)(a^s-e^{-ig})}\right\}$$

$$= (1/2)\left[\frac{e^{ig(t+1)} - 1}{e^{ig} - 1} + \frac{e^{-ig(t+1)} - 1}{e^{-ig} - 1}\right]$$

$$= (1/2)\left[\frac{e^{igt} + e^{-igt} - e^{ig(t+1)} - e^{-ig(t+1)} - e^{ig} - e^{-ig} + 2}{2 - e^{ig} - e^{-ig}}\right]$$

$$= \frac{\cos gt - \cos g(t+1) - \cos g + 1}{2(1 - \cos g)}$$

$$= \frac{\cos gt - \cos g(t+1)}{2(1 - \cos g)} + \frac{1}{2}$$

$$= \frac{\sin g(t + 1/2)}{2 \sin (g/2)} + \frac{1}{2}.$$

Therefore,

$$\sum_{\tau=0}^{t} \cos g\tau = \frac{\sin g(t + 1/2)}{2 \sin (g/2)} + \frac{1}{2}.$$

The sum,

$$1^2 + 1^2 + 2^2 + 1^2 + 2^2 + 3^2 + \cdots + 1^2 + 2^2 + 3^2 + \cdots + t^2,$$

may be determined by (7).

$$\sum_{\tau=0}^{t}\sum_{\sigma=0}^{\tau} \sigma^2 = \mathfrak{D}^{-1}\left\{\frac{a^{2s}}{(a^s - 1)^2} \quad \frac{a^{2s} + a^s}{(a^s - 1)^3}\right\} = \mathfrak{D}^{-1}\left\{\frac{a^{4s} + a^{3s}}{(a^s - 1)^5}\right\}.$$

$$\sum_{\tau=0}^{t}\sum_{\sigma=0}^{\tau} \sigma^2 = \frac{(t+3)^{(4)}}{4!} + \frac{(t+2)^{(4)}}{4!} = \frac{(t+1)(t+2)^{(3)}}{12}.$$

7. Table of transforms. A short list of transforms is given below. See Fort [3] for additional transformations.

	$F(t)$	$f(s)$
1	t^n	$\dfrac{(-1)^n}{(\operatorname{Ln} a)^n} \dfrac{d^n}{ds^n}\left(\dfrac{a^s}{a^s - 1}\right)$
2	$\binom{t+n}{m}$	$\dfrac{a^{(n+1)s}}{(a^s - 1)^{m+1}} - \sum_{j=1}^{n-m} a^{js}\binom{n-j}{m}$, if $m < n$, $\dfrac{a^{(n+1)s}}{(a^s - 1)^{m+1}}$, if $n \leqq m$

3	$\sum_{\tau=0}^{t-1} F(\tau)$	$\frac{1}{a^s - 1} f(s)$
4	$t^n b^{\lambda t}$	$\frac{(-1)^n}{(\text{Ln } a)^n} \frac{d^n}{ds^n}\left(\frac{a^s}{a^s - b^\lambda}\right)$
5	$t^n b^{\lambda t} \sin gt$	$\frac{(-1)^n}{(\text{Ln } a)^n} \frac{d^n}{ds^n}\left(\frac{a^s b^\lambda \sin g}{a^{2s} - 2a^s b^\lambda \cos g + b^{2\lambda}}\right)$
6	$t^n b^{\lambda t} \cos gt$	$\frac{(-1)^n}{(\text{Ln } a)^n} \frac{d^n}{ds^n}\left(\frac{a^{2s} - a^s b^\lambda \cos g}{a^{2s} - 2a^s b^\lambda \cos g + b^{2\lambda}}\right)$
7	$t^n b^{\lambda t} \sinh gt$	$\frac{(-1)^n}{(\text{Ln } a)^n} \frac{d^n}{ds^n}\left(\frac{a^s b^\lambda \sinh g}{a^{2s} - 2a^s b^\lambda \cosh g + b^{2\lambda}}\right)$
8	$t^n b^{\lambda t} \cosh gt$	$\frac{(-1)^n}{(\text{Ln } a)^n} \frac{d^n}{ds^n}\left(\frac{a^{2s} - a^s b^\lambda \cosh g}{a^{2s} - 2a^s b^\lambda \cosh g + b^{2\lambda}}\right)$
9	$t^{(n)} b^{\lambda t} \sinh gt$	$n! a^s b^{n\lambda}\left[\frac{V \cosh ng + U \sinh ng}{(a^{2s} - 2a^s b^\lambda \cosh g + b^{2\lambda})^{n+1}}\right]^*$
10	$t^{(n)} b^{\lambda t} \cosh gt$	$n! a^s b^{n\lambda}\left[\frac{U \cosh ng + V \sinh ng}{(a^{2s} - 2a^s b^\lambda \cosh g + b^{2\lambda})^{n+1}}\right]^*$

* $U+V=(a^s-b^\lambda \cosh g+b^\lambda \sinh g)^{n+1}$, $U-V=(a^s-b^\lambda \cosh g-b^\lambda \sinh g)^{n+1}$.

References

1. E. T. Copson, An Introduction to the Theory of Functions of a Complex Variable, Oxford, 1935.

2. Tomlinson Fort, Infinite Series, Oxford, 1930.

3. ———, Linear difference equations and the Dirichlet series transform, this MONTHLY, vol. 62, 1955, pp. 641–645.

4. ———, Solving difference equations by means of the Dirichlet series transform, Bull. Amer. Math. Soc., vol. 61, 1955, pp. 62–63.

MATHEMATICAL NOTES

EDITED BY IVAN NIVEN, University of Oregon

Because of the large number of papers on hand, consideration of new papers for this department has been temporarily suspended.

THE FUNDAMENTAL THEOREM OF ALGEBRA*

RAYMOND REDHEFFER, University of California, Los Angeles

We prove anew the theorem of Gauss, that a nonconstant polynomial has at least one root. The discussion is based on the Cauchy-Riemann equations,

$$u_x = v_y, \qquad u_y = -v_x, \tag{1}$$

and makes no reference to geometry or topology. The novelty is thought to consist partly in the derivation of (1), but chiefly in the discussion of Remark 4. The latter yields a new proof for the minimum-modulus theorem of function theory.

As background we assume a few elementary facts of differential calculus such as Leibniz' rule and the properties of absolute value. These do not require the completeness of the real-number system, hence can be proved in all rigor in the beginning calculus class. It is necessary to know also that mixed partial derivatives satisfy $u_{xy}=u_{yx}$. Although this is a rather deep theorem in general, it is trivial for polynomials. (Verify the equation directly for $u=a_{pq}x^py^q$; it then follows by linearity for a sum of such terms.)

A second theorem of some complexity is that

$$g_x = g_y = 0, \qquad g_{xx} \geqq 0, \qquad g_{yy} \geqq 0$$

at a minimum of $g(x, y)$. This too is trivial for polynomials. (Without loss of generality let the polynomial $g(x, y)$ have a minimum of value 0 at $x=y=0$. Then

$$g(x, 0) = bx + cx^2 + \cdots + kx^n = x(b + cx + \cdots + kx^{n-1}),$$

where $b, c, \cdots, k$ are constant. The assumption that $b\neq 0$ shows immediately that $x=0$ is not a minimum. When $b=0$ the same procedure shows that $c\geqq 0$).

The only other "deep" theorem we need is that a continuous function in the closed circle attains a minimum in the closed circle. This is not trivial even for polynomials; in fact the fundamental theorem of algebra is deducible therefrom.

1. Algebraic properties of the Cauchy-Riemann equations. We say that a complex function $w(x, y)$ satisfies the Cauchy-Riemann equations if its real and imaginary parts do so; that is, if $w(x, y)=u(x, y)+iv(x, y)$ where u and v are real and satisfy (1).

* The preparation of this paper was sponsored jointly by the Office of Naval Research and the Office of Ordnance Research.

Remark 1. If w and W are complex functions which satisfy the Cauchy-Riemann equations, then so are $w+W$ and wW.

If $w=u+iv$, $W=U+iV$, then $w+W=(u+U)+i(v+V)$. Now the equations (1) for w and W yield

$$(u+U)_x = u_x + U_x = v_y + V_y = (v+V)_y$$

which is the first equation for $w+W$. The second follows similarly.

Again, we have $wW=uU-vV+i(uV+vU)$. By Leibniz' rule,

$$(uU-vV)_x = uU_x + u_xU - vV_x - v_xV,$$
$$(uV+vU)_y = uV_y + Vu_y + vU_y + Uv_y.$$

That these expressions are equal follows by comparing coefficients of u, U, v, V, with due regard to (1). We now have the first equation for wW; proof of the second is similar.

Remark 2. Let u and v be real and imaginary parts of a polynomial, so that

$$u+iv = a_nz^n + \cdots + a_1z + a_0,$$

where $z=x+iy$ and the a's are constant. Then the polynomials u and v satisfy (1).

The reader may verify that $x+iy$ satisfies the Cauchy-Riemann equations; that is, $u=x$, $v=y$ satisfy (1). Hence, by Remark 1, the same is true of $(x+iy)(x+iy)=(x+iy)^2$ and of $(x+iy)(x+iy)^2=(x+iy)^3$, and so on. Hence, $(x+iy)^m$ satisfies the Cauchy-Riemann equations for each integer m.

Again, a constant $a+ib$ is clearly a solution; that is, $u=a$, $v=b$ satisfies (1). Hence, by Remark 1, $a_m(x+iy)^m$ satisfies the Cauchy-Riemann equations; and, again by Remark 1, a sum of such terms does so.

2. A minimum modulus theorem. In the following remarks, u and v are solutions of (1) as given by Remark 2.

Remark 3. If $2f=u^2+v^2$, and if $m=u_x^2+u_y^2$, then

$$f_x^2 + f_y^2 = m(u^2+v^2), \tag{2}$$

$$f_{xx} + f_{yy} = 2m. \tag{3}$$

For proof, we use (1) to obtain three expressions for f_x and f_y:

$$f_x = uu_x + vv_x = uu_x - vu_y = uv_y - vu_y,$$
$$f_y = uu_y + vv_y = uu_y + vu_x = -uv_x + vu_x.$$

The second of these expressions gives

$$f_x^2 + f_y^2 = (uu_x - vu_y)^2 + (uu_y + vu_x)^2 = m(u^2+v^2)$$

as desired. The third of these expressions gives

$$f_{xx} = uv_{yx} + u_x v_y - vu_{yx} - v_x u_y,$$
$$f_{yy} = -uv_{xy} - u_y v_x + vu_{xy} + v_y u_x,$$

whence, by addition and by (1),

$$f_{xx} + f_{yy} = u_x v_y - u_y v_x - v_x u_y + v_y u_x$$
$$= u_x^2 + u_y^2 + u_y^2 + u_x^2 = 2m.$$

Remark 4. With $f(x, y)$ as above, suppose $f(0, 0) = c$, and suppose $f(x, y) \geqq c+1$ on the circle $x^2+y^2=r^2$. Then $f=0$ at some point inside this circle.

For proof, let h be a small positive constant and set

$$g(x, y) = f(x, y) - hx^2. \tag{4}$$

At the origin and on the circle $x^2+y^2=r^2$ we have, respectively,

$$g(0, 0) = c, \tag{5}$$

$$g(x, y) \geqq c + 1 - hx^2 \geqq c + 1 - hr^2 > c, \tag{6}$$

provided h is sufficiently small.

The function $g(x, y)$ attains an absolute minimum in $x^2+y^2 \leqq r^2$ since g is continuous. Moreover, this minimum is not attained on the boundary, since (5) and (6) show that $g(0, 0)$ is smaller than any value attained on the boundary. There is, then, an interior minimum.

At the minimum point we have, by (4),

$$f_x - 2hx = g_x = 0, \qquad f_y = g_y = 0; \tag{7}$$

$$f_{xx} - 2h = g_{xx} \geqq 0, \qquad f_{yy} = g_{yy} \geqq 0. \tag{8}$$

Equations (7) combine with (2), and (8) with (3) to give, respectively,

$$m(u^2 + v^2) = f_x^2 + f_y^2 = 4h^2x^2, \tag{9}$$

$$2m = f_{xx} + f_{yy} \geqq 2h. \tag{10}$$

By division of (9) and (10), $(u^2+v^2)/2 \leqq 2hx^2 \leqq 2hr^2$. Since h is arbitrary this shows that u^2+v^2 cannot have a positive lower bound in the circle, whence we deduce $u^2+v^2=0$ at some point.

3. The fundamental theorem. Remark 4 contains the fundamental theorem of algebra as a special case. For, if

$$u + iv = a_n z^n + \cdots + a_1 z + a_0, \tag{11}$$

where $z=x+iy$, $a_n \neq 0$, and $n \geqq 1$, then we have the Cauchy-Riemann equations

(1), by Remark 2. And we also have (trivially)

$$\lim \frac{u+iv}{a_n z^n} = 1 \qquad \text{as } |z| \to \infty.$$

Hence, for large $|z|$ we have $|u+iv|/|a_n z^n| > 1/2$ say, or, equivalently, $|u+iv| > (1/2)|a_n||z|^n$. This shows that $u^2+v^2 \to \infty$ as $x^2+y^2 \to \infty$, and therefore that the hypothesis of Remark 4 is satisfied if r is large enough. The point (x, y) where u and v both vanish yields a complex number $z=x+iy$ where the polynomial vanishes, and the existence of a root is established.

NOTE ON THE DISTRIBUTION OF THE TOTATIVES

PAUL J. MCCARTHY, Florida State University

If n is a positive integer, the $\phi(n)$ numbers $\leqq n$ and relatively prime to n are called the *totatives* of n.* We shall be concerned in this note with the distribution of the totatives of n in certain subintervals from 0 to n. In fact, the intervals are given by

$$nq/k < m < n(q+1)/k, \qquad q = 0, 1, \cdots, k-1. \tag{1}$$

No totative can occur as the endpoint of such an interval if we make the natural assumption that $n > k$. We make this assumption throughout this note.

We denote by $\phi(k, q, n)$ the number of totatives of n in the interval (1). Then a natural measure of the distribution of the totatives of n in the intervals (1) is the function

$$E(k, q, n) = \phi(n) - k\phi(k, q, n). \tag{2}$$

The totatives of n are said to be *uniformly distributed* with respect to k if $E(k, q, n) = 0$ for all q: then each of the intervals (1) contains $\phi(n)/k$ of the totatives of n. Lehmer* called n *exceptional* with respect to k if n is divisible by k^2 or by a prime of the form $kx+1$. He showed that the totatives of n are uniformly distributed with respect to k when n is exceptional with respect to k. Thus we have a sufficient condition for the uniform distribution of the totatives of n. It is not a necessary condition as can be seen from the fact that the totatives 1, 2, 4, 5, 7, 8 of 9 are uniformly distributed with respect to 6. In fact, we have the following result.

THEOREM. *If k is a prime, then the totatives of n are uniformly distributed with respect to k, if and only if n is exceptional with respect to k. Moreover, if k is not squarefree there is an integer $n > k$ such that the totatives of n are uniformly distributed with respect to k, but n is not exceptional with respect to k.*

We need only prove the necessity of the condition. From (2), $E(k, q, n) \equiv \phi(n) \pmod{k}$, and since $\phi(n)$ is multiplicative we have for $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_t^{\alpha_t}$,

* D. H. Lehmer, The distribution of totatives, Canad. J. Math., vol. 7, 1955, pp. 347–357.

$$E(k, q, n) \equiv \prod_{i=1}^{t} p_i^{\alpha_i - 1}(p_i - 1) \pmod{k}.$$

Suppose $E(k, q, n) = 0$. Then, since k is prime, either $p_i \equiv 1 \pmod{k}$ for some i, or $k = p_i$ and $\alpha_i > 1$ for some i. In either case, n is exceptional with respect to k.

To prove the second part of the theorem we shall make use of the

LEMMA. *If k divides n, and n/k has the same prime divisors as n, then the totatives of n are uniformly distributed with respect to k.*

If $r < n/k$ is prime to n, so also is $r + qn/k$ prime to n for all q, and *vice versa*. Hence the number of totatives of n in the interval (1) is independent of q, so all the $E(k, q, n)$ are equal, and by (6) of Lehmer's paper* each $E(k, q, n) = 0$.

Now suppose k is not squarefree and write $k = p^{\alpha}k_1$, where $\alpha > 1$ and k_1 is prime to p. Let $n = p^{\alpha+1}k_1^2$. By the lemma the totatives of n are uniformly distributed with respect to k. Clearly k^2 does not divide n. If n has a prime divisor of the form $kx + 1$, then so does k, which is impossible. Thus n is not exceptional with respect to k. This completes the proof of the theorem.

Remark. The case when k is squarefree and composite presents added difficulties. The example preceding the statement of the theorem shows that the first part of the theorem is not true for such k (see also Theorem 7 of Lehmer's paper*), and we have been unable to determine whether or not the second part is true.

We wish to thank the referee for his kind suggestions.

CLASSROOM NOTES

EDITED BY C. O. OAKLEY, Haverford College

All material for this department should be sent to C. O. Oakley, Department of Mathematics, Haverford College, Haverford, Pa.

PROBABILITY AND SUMS OF SERIES

G. B. THOMAS, JR., Massachusetts Institute of Technology and Stanford University

Many interesting problems in probability stem from the simple model of repeated Bernoulli trials, where the probability of success is the same at each trial. If the trials are independent, but the probability p_n of success at trial n is allowed to vary with n, the trials are called Poisson trials [1, p. 233]). When we consider an infinite sequence of Poisson trials we are led quite naturally to certain associated series (for example, generating functions, *cf.* Feller, Chapter 11). Here we consider the probability of eventually obtaining a success. This is given by

$$(1) \qquad Pr \text{ (success on some trial)} = \sum_{n=1}^{\infty} f_n,$$

where f_n is the probability of *first* success at trial n,

$$(2) \qquad f_1 = p_1, \qquad f_n = (1 - p_1)(1 - p_2) \cdots (1 - p_{n-1})p_n, \qquad n > 1.$$

In particular, we are interested in circumstances under which the probability of getting at least one success is one. The following theorem gives the answer. By specializing the p_n we also obtain the sums of certain infinite series in explicit form.

THEOREM. *Let* $\{p_n\}$ *be a sequence of real numbers,* $0 \leqq p_n < 1$. *Let* f_n *be given by* (2). *Then*

$$(3) \qquad \sum_{n=1}^{\infty} f_n \leqq 1,$$

and equality holds in (3), *if and only if*

$$(4) \qquad \prod_{n=1}^{\infty} (1 - p_n) \text{ } \textit{diverges to zero,}$$

or (*equivalently*), *if and only if*

$$(5) \qquad \sum_{n=1}^{\infty} p_n \text{ } \textit{diverges to plus infinity.}$$

Proof. Let N be a positive integer. Let

$$(6) \qquad P_N = \sum_{1}^{N} f_n, \qquad Q_N = \prod_{1}^{N} (1 - p_n).$$

Then P_N is the probability of at least one success in the first N trials, while Q_N is the probability of no successes in the same N trials. Hence

$$(7) \qquad P_N + Q_N \equiv 1.$$

Equation (7) is an identity for all sequences $\{p_n\}$ and for all positive integers N. In case $0 \leqq p_n \leqq 1$, the sequence $\{Q_N\}$ is monotone nonincreasing and bounded below by zero. Hence Q_N tends to a limit as N increases indefinitely. Likewise P_N has a limit. We denote these limits by Q and P respectively and have from (7)

$$(8) \qquad P + Q = 1,$$

identically for all sequences $\{p_n\}$ such that $0 \leqq p_n \leqq 1$, with

$$(9) \qquad P = \sum_{1}^{\infty} f_n \leqq 1, \qquad Q = \prod_{1}^{\infty} (1 - p_n) \geqq 0.$$

From (8) it follows at once that $P = 1$ if and only if $Q = 0$. And $Q = 0$ if and only

if either (i) $p_n=1$ for some n, or (ii) $\sum p_n$ diverges to plus infinity [2]. The theorem is thus established.

Example 1. Let $p_n=1/(n+1)$ and obtain

$$\sum_1^\infty \frac{1}{n(n+1)} = 1.$$

Example 2. Let $p_n=n/(n+1)$ and obtain

$$\sum_1^\infty \frac{n}{(n+1)!} = 1.$$

Example 3. Let $p_n=n/(n+k)$, $k>0$, and obtain

$$\sum_{n=1}^\infty \frac{nk^{n-1}}{\Gamma(n+k+1)} = \frac{1}{\Gamma(k+1)}.$$

In particular, for $k=1/2$,

$$\sum_{n=1}^\infty n \Big/ \left(\prod_{k=1}^n (2k+1)\right) = 1/2.$$

Remarks. (a) The identity (7) may be of interest to algebra teachers looking for another easy exercise in mathematical induction. Feller's book provides many more. Their probabilistic interpretations may provide additional interest.

(b) The series in Example 2 above is clearly closely related to the series for e^x. In fact, if we look at the generating function for the sequence $\{f_n\}$ we have

$$F(x) \equiv \sum_1^\infty f_n x^n = \sum_1^\infty nx^n/(n+1)!,$$

and we soon recognize this as being x times the derivative of

$$G(x) = \sum_1^\infty x^n/(n+1)! = x^{-1}(e^x - 1 - x).$$

Hence

$$F(x) = x\frac{d}{dx}G(x) = e^x - x^{-1}(e^x - 1),$$

and the series in Example 2 is evaluated by taking $x=1$; $F(1)=1$. To obtain the expected number of trials until first success occurs, we would compute

$$F'(1) = \sum_{n=1}^\infty nf_n = e - 1.$$

(c) One of the Borel-Cantelli lemmas [1, p. 155] asserts: *If the events A_k are*

mutually independent, and if $\sum Pr(A_k)$ *diverges, then with probability one infinitely many* A_k *occur.*

We may take A_k to be the event, success at trial k, and apply this lemma and our theorem to obtain the following interesting corollary.

COROLLARY. *Let* p_n *be the probability of success on the nth trial in a sequence of Poisson trials. If* $0 \leq p_n < 1$, *and the probability of obtaining at least one success is one, then the probability of obtaining infinitely many successes is also one.*

References

1. W. Feller, An Introduction to Probability and its Applications, New York, 1950.
2. J. M. Hyslop, Infinite Series, New York, 1942, p. 94.

THE GEOMETRY OF VARIATION OF PARAMETERS

R. M. CONKLING, New Mexico College of Agriculture and Mechanic Arts

The usual first course in differential equations treats the method of variation of parameters for finding particular integrals of nth order linear nonhomogeneous differential equations somewhat formally. The method of Lagrange multipliers for extreme values of functions of several variables under constraints often suffers from the same kind of treatment. There is, of course, a geometric interpretation of both methods (see the treatment of Lagrange multipliers in Kaplan, *Advanced Calculus*, Cambridge, 1952). Since the variation of parameters method is, for most students, the first encounter with perturbation theory, it seems appropriate to provide at least the better students with some geometry to go with it. It is the purpose of this note to present a rather simple example together with suggestions as to what happens in the general case.

We consider the second order nonhomogeneous differential equation with initial conditions

$$\ddot{x} + x = \mu f(x, \dot{x}, t), \qquad x(0) = u, \qquad \dot{x}(0) = v.$$

The homogeneous equation is that of a harmonic oscillator with angular velocity unity, whose solution is $x(t) = u \cos t + v \sin t$. The variation of parameters method seeks to find $u(t)$ and $v(t)$ so that $u(t) \cos t + v(t) \sin t$ is a solution of the nonhomogeneous (perturbed) equation.

Changing (1) to the corresponding first order system, we have

$$(1) \qquad \begin{aligned} \dot{x} &= y, & x(0) &= u, \\ \dot{y} &= -x + f(x, y, t), & y(0) &= v, \end{aligned}$$

and the unperturbed system

$$(2) \qquad \begin{aligned} \dot{x} &= y, & x(0) &= u, \\ \dot{y} &= -x, & y(0) &= v. \end{aligned}$$

The solutions in the xy-(phase) plane of the conservative (homogeneous) system

(2) are given parametrically by

$$x = u \cos t + v \sin t = \phi(u, v, t),$$
$$y = -u \sin t + v \cos t = \psi(u, v, t). \tag{3}$$

Eliminating t, we have

$$\frac{dy}{dt} \Big/ \frac{dx}{dt} = \frac{dy}{dx} = -\frac{x}{y},$$

so that the solutions are concentric circles. The representative point travels clockwise on one of these circles (determined by the initial values u and v) with angular velocity unity.

Equations (3), considered as the equations of a rotation through the angle t, transform the xy system of coordinates into a system of coordinates which rotates clockwise with angular velocity unity. In the new coordinate system, the point (u, v) is an equilibrium state for the system (2). Regarding (3), then, as a change of variables, and substituting into (1), we have

$$\frac{\partial \phi}{\partial u}\dot{u} + \frac{\partial \phi}{\partial v}\dot{v} + \frac{\partial \phi}{\partial t} = \psi(u, v, t), \tag{4}$$

$$\frac{\partial \psi}{\partial u}\dot{u} + \frac{\partial \psi}{\partial v}\dot{v} + \frac{\partial \psi}{\partial t} = -\phi(u, v, t) + \mu f(\phi, \psi, t). \tag{5}$$

On the other hand, regarding (3) as the solutions of (2), we have

$$\frac{d\phi}{dt} \overset{t}{\equiv} \psi(u, v, t), \tag{6}$$

$$\frac{d\psi}{dt} \overset{t}{\equiv} -\phi(u, v, t), \tag{7}$$

$$\phi(0) = u, \qquad \psi(0) = v. \tag{8}$$

By virtue of (6), (4) becomes

$$\frac{\partial \phi}{\partial u}\dot{u} + \frac{\partial \phi}{\partial v}\dot{v} = 0.$$

This equation is arrived at in the usual treatment by imposing one of two allowable conditions on the functions $u(t)$ and $v(t)$, the other being the result of using (7) in (5). In the harmonic oscillator example, where, if u and v are regarded as unknown functions of t in (3), we have $x = -u \sin t + v \cos t + \dot{u} \cos t + \dot{v} \sin t$, and one of the conditions imposed on u and v is that $\dot{u} \cos t + \dot{v} \sin t = 0$. It is implied that any condition could be imposed at this point, but it is clear that in this geometric setting, the usual condition is the only reasonable one. Using (6) and (7) to simplify (4) and (5),

$$\frac{\partial \phi}{\partial u}\dot{u} + \frac{\partial \phi}{\partial v}\dot{v} = 0, \quad \frac{\partial \psi}{\partial u}\dot{u} + \frac{\partial \psi}{\partial v}\dot{v} = \mu f(\phi, \psi, t). \tag{9}$$

Since, when $t=0$, the determinant of the coefficients (the Jacobian J of the transformation (3)) is unity, we can solve for u and v in some neighborhood of $t=0$:

$$\dot{u} = -\frac{\mu}{J} f(\phi, \psi, t)\frac{\partial \phi}{\partial v}, \qquad \dot{v} = \frac{\mu}{J} f(\phi, \psi, t)\frac{\partial \phi}{\partial u}. \tag{10}$$

In the harmonic oscillator problem, these become $\dot{u} = -\mu f(\phi, \psi, t) \sin t$, $\dot{v} = \mu f(\phi, \psi, t) \cos t$.

In any case, u and v will be linear in $\mu f(\phi, \psi, t)$, the perturbing function, so that for small values of the constant μ, the rates of change of u and v will be small, *i.e.*, u and v are nearly constant. The quantities $u(t)-u(0)$, $v(t)-v(0)$ are perturbations, measuring the change that the initial point of the solution of the homogeneous equation undergoes as a result of the perturbing function $\mu f(x, \dot{x}, t)$.

If, for the harmonic oscillator, we have $\mu f(x, \dot{x}, t) = \mu e^{-t}$, the perturbations are

$$\begin{aligned} u(t) - u(0) &= -\mu\left[\frac{e^{-t}}{2}(-\sin t - \cos t) + 1/2\right], \\ v(t) - v(0) &= \mu\left[\frac{e^{-t}}{2}(-\cos t + \sin t) + 1/2\right]. \end{aligned} \tag{11}$$

For large t, $u(t) \approx u(0) - \mu/2$, and $v(t) \approx v(0) + \mu/2$, so that the steady state solution of the perturbed equation in the phase plane is the circle whose center is at the origin, passing through $(u-\mu/2, v+\mu/2)$. It is interesting to note that there finally results no change in the orbit if the initial conditions and the parameter μ are such that $u-v=\mu/2$, no matter how large is μ. The perturbing function μe^{-t} produces no net change in the energy state of the oscillator in this case. However, in general, small perturbations are associated only with small values of μ and short intervals of time. The steady state solution may be greatly different from the unperturbed solution, since the perturbing function has had a long time in which to change the energy level of the system, even though this level changes very slowly when μ is small.

In the above example, the perturbations, viewed in the uv-plane, appear as the curves given parametrically by equations (11). That is, if one were riding along on the trajectory of the solution of the unperturbed equation, one would view only the effect of the perturbing function.

For the relation of the variation of parameters method to one-sided Green's functions, see Miller, *Engineering Mathematics*, New York, 1956.

References

1. A. S. Andronow and C. E. Chaikin, Theory of Oscillations, Princeton, 1949.
2. F. R. Moulton, Celestial Mechanics, New York, 1914.

ELEMENTARY PROBLEMS AND SOLUTIONS

EDITED BY HOWARD EVES, University of Maine

Send all communications concerning Elementary Problems and Solutions to Howard Eves, Mathematics Department, University of Maine, Orono, Maine. This department welcomes problems believed to be new, and demanding no tools beyond those ordinarily furnished in the first two years of college mathematics. To facilitate their consideration, solutions should be submitted on separate, signed sheets, within three months after publication of problems.

PROBLEMS FOR SOLUTION

E 1281. *Proposed by D. J. Newman, AVCO Research Division, Lawrence, Mass.*

Prove that no perfect square is 7 more than a perfect cube.

E 1282. *Proposed by Leon Bankoff, Los Angeles, Calif.*

Tangents PQ, PR are drawn from a point P to a circle (O). Another tangent touches the circle at a point B on the minor arc QR and cuts PQ, PR at A, C. Show that QR is equal to the sum of the sides of the triangle of minimum perimeter that can be inscribed in triangle OCA.

E 1283. *Proposed by J. W. Andrushkiw, Seton Hall University*

If three consecutive coefficients of the real polynomial

$$P(x) = x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n$$

are equal and different from zero, show that the zeros of $P(x)$ are not all real.

E 1284. *Proposed by Charles Fox, McGill University*

Let P be any point on a central conic C, and let the normal to C at P and the polar of P with respect to the director circle of C meet at Q. Show that PQ is equal to the radius of curvature of C at P. What form does this result take for a parabola?

E 1285. *Proposed by James Ax and Lawrence Shepp, Polytechnic Institute of Brooklyn*

Let A and B be two nth order determinants, and let C be a third determinant whose (i, j)th element is the determinant A with its ith column replaced by the jth column of B. Show that $C = A^{n-1}B$.

SOLUTIONS

Willie's Age

E 1251 [1957, 109]. *Proposed by J. S. Lew, Princeton University*

"Did your teacher give you that problem?" I asked. "It looks rather tedious." "No," said Willie, "I made it up. It's a polynomial with my age as a root.

That is, x stands for my age last birthday."

"Well, then," I remarked, "It shouldn't be so hard to work out—integer coefficients, integral root. Suppose I try $x=7 \cdots$. No, that gives 77."

"Do I look only seven years old?" demanded Willie.

"Well, let me try a larger integer $\cdots$. No, that gives 85, not zero."

"Oh, stop kidding!" said Willie, looking over my shoulder. "You know I'm older than that!"

How old is Willie?

Solution by D. C. B. Marsh, Colorado School of Mines. Willie's friend, dispensing with trial-and-error, might make use of the fact that $a-b$ is an integral divisor of $P(a)-P(b)$ when a, b are distinct integers and $P(x)$ is a polynomial with integral coefficients. Denoting the "larger integer" tried by I and Willie's age by A, we have $I-7$ divides 8, $A-7$ divides 77, $A-I$ divides 85, and $7<I<A$. It follows that I must be one of 8, 9, 11, 15 and A one of 14, 18, 84. To have $A-I$ divide 85, we see that the second integer tried was 9 and that Willie is 14 years old.

Also solved by Merrill Barnebey, H. F. Bennett, J. L. Botsford, C. K. Bradshaw, Julian Braun, D. A. Breault, W. B. Carver, J. E. Davis, Underwood Dudley, P. L. Duren, Hazel E. Evans, M. A. Feldstein, H. J. Fletcher, L. R. Ford, C. B. Germain, Cornelius Groenewoud, A. R. Hyde, I. M. Isaacs, Ray Jurgensen, Seymour Kass, Andrew Kraus, Joe Lipman, Wallace Manheimer, J. B. Muskat, C. S. Ogilvy, W. D. Peeples, Jr., C. F. Pinzka, R. I. Purry, L. A. Ringenberg, David Rosen, Azriel Rosenfeld, David Rothman, D. K. Stillinger, Chih-yi Wang, R. M. Warten, Alan Wayne, and the proposer. Late solution by L. D. Goldstone.

Carver, Ford, and Wayne pointed out that Willie's polynomial must be of the form

$$(x-7)(x-9)(x-14)Q(x)-3x^2+52x-140,$$

where $Q(x)$ is any polynomial with integral coefficients.

A Property of Positive Integral Powers of an Integer

E 1252 [1957, 109]. *Proposed by Jerome Rothstein, Signal Corps Engineering Laboratories, Fort Monmouth, N. J.*

Prove that the difference between two positive integral powers of the same integer is exactly divisible by six unless the integer gives the remainder two on division by three and one power is odd while the other is even. Show that in this exceptional situation the sum of the two powers is exactly divisible by six.

Solution by Andrew Kraus, University of Colorado. Let n, m be positive integers, and a an integer. In every case $a^n \pm a^m \equiv 0 \pmod 2$. If $a \equiv 0 \pmod 3$, then $a^n-a^m=(3k)^n-(3k)^m \equiv 0 \pmod 3 \equiv 0 \pmod 6$. If $a \equiv 1 \pmod 3$ and n, m are of like parity, then $a^n-a^m=(3k+1)^n-(3k+1)^m \equiv 0 \pmod 3 \equiv 0 \pmod 6$. If $a \equiv -1 \pmod 3$, and n, m are of unlike parity, then $a^n+a^m=(3k-1)^n+(3k-1)^m \equiv 0 \pmod 3 \equiv 0 \pmod 6$.

Also solved by J. L. Alperin, R. V. Andree, Leon Bankoff, H. F. Bennett, P. L. Chessin, J. E. Darraugh, A. G. Davis, Irma Esrig, Hazel E. Evans, Michael Goldberg, Cornelius Groenewoud, Emil Grosswald, J. H. Hodges, J. T. Humphrey, A. R. Hyde, I. M. Isaacs, J. B. Johnson, Ray Jurgensen, Sidney Kravitz, Aaron Lieberman, Joe Lipman, Marshall Luban, W. S. McCulley, Wallace Manheimer, D. C. B. Marsh, L. C. Marshall, G. E. Meador, David Muskat, J. B. Muskat,

C. S. Ogilvy, Hiram Paley, D. S. Passman, D. J. Persico, C. F. Pinzka, L. A. Ringenberg, Azriel Rosenfeld, Kenneth Ross, David Rothman, Frank Saunders, J. A. Schumaker, Lawrence Shepp, W. A. Soper, Jr., D. R. Sudborough, Eldon Vought, Chih-yi Wang, R. M. Warten, Maud Willey, L. K. Williams, and the proposer. Late solution by D. A. Breault.

Application of a Formula of Finite Differences

E 1253 [1957, 109]. *Proposed by T. R. Jenkins, University of Idaho, and by Nathaniel Macon and Abraham Spitzbart, Flight Propulsion Laboratory, General Electric Co.*

For n and p positive integers with $n>p$, and x and y arbitrary, show that

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} (x + ky)^p = 0.$$

Find the value of the sum for $n=p$.

I. *Solution by Chih-yi Wang, University of Minnesota.* We have

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} (x + ky)^p = \frac{d^p}{dt^p} \left\{ \sum_{k=0}^{n} (-1)^k \binom{n}{k} e^{(x+ky)t} \right\} \Big|_{t=0}$$

$$= \frac{d^p}{dt^p} \left\{ e^{xt}(1 - e^{yt})^n \right\} \Big|_{t=0}.$$

For $n>p$, the pth derivative of the above expression involves all terms containing the factors $(1-e^{yt})^r$, $n-p \leqq r \leqq n$, whence the stated result follows. For $n=p$, the only nonvanishing term is $n!(-y)^n$.

II. *Solution by W. A. Al-Salam, Duke University.* Denote the sum by S_p. Expanding $(x+ky)^p$ and changing order of summation, we find

$$S_p = \sum_{s=0}^{p} \binom{p}{s} x^{p-s} y^s \sum_{k=0}^{n} (-1)^k \binom{n}{k} k^s.$$

But the inside sum vanishes for $s<n$ and has the value $(-1)^n n!$ if $s=n$. Hence $S_p=0$ or $(-1)^n y^n n!$ according as $p<n$ or $p=n$.

III. *Solution by R. S. Pinkham, Princeton University.* From the calculus of finite differences

$$\Delta^n w^p = \sum_{k=0}^{n} (-1)^{n+k} \binom{n}{k} (w + k)^p.$$

Setting $w=x/y$ and multiplying both sides by $(-1)^n y^p$, we find

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} (x + ky)^p = \Delta^n w^p (-1)^n y^p.$$

But when $n \geqq p$, $\Delta^n w^p = \delta_p^n n!$, and we obtain the desired results.

Also solved by J. L. Alperin, J. L. Brown, Jr., J. E. Darraugh and F. D. Parker (jointly), R. V. Esperti, C. B. Germain, Cornelius Groenewoud, Emil Grosswald, R. H. Hou, J. B. Johnson, P. G.

Kirmser, M. S. Klamkin, Joseph Lehner, Aaron Lieberman, Joe Lipman, D. C. B. Marsh, Mary P. Moseley, C. F. Pinzka, Ronald Pyke, L. A. Ringenberg, J. B. Roberts, D. A Robinson, E. P. Rozycki, Paul Schaefer, F. C. Smith, E. P. Starke, Julius Vogel, David Zeitlin, and the proposers.

Pinzka pointed out that the desired conclusion follows from Problem 4183 [1947, 235], and that Problem 3625 [1934, 454] also deals with the limiting case. For the valuation of the inside sum in Solution II see, *e.g.*, Schwatt, *An Introduction to Operations with Series*, p. 101, or Chrystal, *Algebra*, Part II, 2nd ed., formulas (3) and (4) p. 207 (by setting $x = -n$ and introducing the factor $(-1)^n$). The fundamental formula of the calculus of finite differences used in Solution III may be found in Jordan, *Calculus of Finite Differences*, p. 132 *et seq.*, or Boole, *Calculus of Finite Differences*, 3rd ed., formula (3), p. 19.

Pencils of Conics Containing a Circle

E 1254 [1957, 109]. *Proposed by Robin Robinson, Dartmouth College*

Prove that if two conics intersect in four distinct points, these points are concyclic if and only if the axes of the two conics are parallel or perpendicular.

I. *Solution by L. R. Ford, Illinois Institute of Technology.* Let the conics be the noncircles

$$S_i = A_i x^2 + H_i xy + B_i y^2 + E_i x + F_i y + C_i = 0, \qquad i = 1, 2.$$

The general equation of a conic through their points of intersection is

$$S_1 + kS_2 = (A_1 + kA_2)x^2 + (H_1 + kH_2)xy + (B_1 + kB_2)y^2 + \cdots = 0.$$

This is made a circle by taking $A_1+kA_2=B_1+kB_2$, $H_1+kH_2=0$, which is possible if and only if $H_1/(A_1-B_1)=H_2/(A_2-B_2)$. This can be written $\tan 2\theta_1 = \tan 2\theta_2$, where θ_i is the angle between the x-axis and an axis of the conic. Hence, $2\theta_1=2\theta_2+n\pi$, or $\theta_1=\theta_2+n\pi/2$, as required.

II. *Solution by Roland Deaux, Faculté Polytechnique de Mons, Belgium.* By virtue of Desargues' theorem, conics passing through the vertices A, B, C, D of a proper quadrangle determine on the line at infinity pairs of points of an involution I. In order that A, B, C, D be concyclic, it is necessary and sufficient that the circular points U, V constitute a pair of I. The projection of I from any proper point P is an involution i which, because of the pair formed by the isotropic lines PU, PV, has two perpendicular double rays p_1, p_2, and these lines are thus the bisectors of the angles formed by any pair of conjugate lines belonging to i. The pencil of conics considered contains two parabolas whose axes are parallel to p_1, p_2, and these lines are parallel to the axes of any other conic of the pencil, whence the property.

Also solved by W. A. Al-Salam, N. A. Court, Michael Goldberg, A. R. Hyde, Joe Lipman, Wallace Manheimer, D. C. B. Marsh, Beckham Martin, C. S. Ogilvy, O. J. Ramler, Sister M. Stephanie, Chih-yi Wang, Maud Willey, and an anonymous solver.

R. H. F. Denniston gave the references: Durell, *Algebraic Geometry*, London, 1955, p. 351; Robson, *An Introduction to Analytic Geometry*, vol. ii, Cambridge, 1947, p. 97; Walker, *Cartesian and Projective Geometry*, London, 1953, pp. 193 (analytic proof), 276 (projective proof). Al-Salam pointed out that the desired result follows immediately from the theorem, "The common chords of a conic and a circle taken in pairs are equally inclined to the axis of the conic," and he gave the reference, Askwith, *Analytic Geometry of the Conic Sections*, London, 1950, p. 218.

ADVANCED PROBLEMS AND SOLUTIONS

EDITED BY E. P. STARKE, Rutgers University

Send all communications concerning Advanced Problems and Solutions to E. P. Starke, Rutgers University, New Brunswick, New Jersey. All manuscripts should be typewritten with double spacing and margins at least one inch wide. Problems containing results believed to be new or extensions of old results are especially sought. Proposers of problems should also enclose any solutions or information that will assist the editor. In general, problems in well-known textbooks or results in readily accessible sources should not be proposed for this department.

PROBLEMS FOR SOLUTION

4753. *Proposed by J. K. Senior, University of Chicago*

Are there any integers larger than 31 which can be represented in more than one way by the form $(a^m-1)/(a-1)$, a and m integers and $m>2$.

4754. *Proposed by T. G. Room, University of Sydney, Australia*

Given the four quadrics

$$L \equiv x^2 + dyz - cyt + bzt = 0, \qquad M \equiv y^2 + dzx + cxt - azt = 0,$$
$$N \equiv z^2 + dxy - bxt + ayt = 0, \qquad K \equiv t^2 - ayz - bzx - cxy = 0.$$

Show that if $1+a^3+b^3+c^3+d^3+3abcd=0$ then they have six points in common, and otherwise, none.

4755. *Proposed by Chandler Davis, Columbia University*

In how many ways can the first n positive integers be arranged in alternately increasing and decreasing order? That is, how many permutations $\pi: \pi(1), \cdots, \pi(n)$ are there such that the quantities $(-1)^k\{\pi(k+1)-\pi(k)\}$, for $k=1, \cdots, n-1$, have all the same sign?

4756. *Proposed by J. L. Massera, Institute of Mathematics and Statistics, Montevideo, Uruguay*

Let $p(x_1, x_2, \cdots, x_n)$, $q(x_1, x_2, \cdots, x_n)$ be two real functions of n real variables x_i, defined and continuous in a parallelotope $R: 0 \leqq x_i \leqq a_i < \infty$. Assume that $p(x_1, \cdots)=q(x_1, \cdots)=0$ whenever $x_1x_2 \cdots x_n=0$, and that $p(x_1, \cdots) > 0$, $q(x_1, \cdots) \geqq 0$, when $x_1x_2 \cdots x_n \neq 0$. Prove that there exists a real function $h(u)$ of a real variable u, defined, continuous and strictly increasing for $u \geqq 0$, $h(0)=0$, such that throughout R

$$h[q(x_1, \cdots)] < p(x_1, \cdots).$$

4757. *Proposed by O. P. Aggarwal, University of Washington, Seattle*

Prove for every integer $n \geqq 0$, and for any positive c,

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} \frac{1}{(c+k)^2} = \frac{\Gamma(c)\Gamma(n+1)}{\Gamma(c+n+1)} \sum_{k=0}^{n} \frac{1}{c+k}.$$

SOLUTIONS

Convergence of a Sequence

4703 [1956, 584]. *Proposed by R. E. Bellman, Rand Corporation, Santa Monica, California*

Show that the conditions (a) $u_n \geqq u_{n-1} - a_{n-1}$, $a_n \geqq 0$, $\sum_{n=0}^{\infty} a_n < \infty$, (b) $u_n \leqq M$ imply the convergence of the sequence $\{u_n\}$.

I. *Solution by E. M. Wright, King's College, Aberdeen, Scotland.* Put $A_n = \sum_{m=1}^{n} a_m \to A$ as $n \to \infty$. Then

$$u_{n-1} + A_{n-2} \leqq u_n + A_{n-1} < M + A,$$

and so $\{u_n + A_{n-1}\}$ is a convergent sequence. Hence, so is $\{u_n\}$.

II. *Solution by R. P. Agnew, Cornell University.* The hypothesis can be weakened to (a′) $u_n \geqq u_{n-1} - a_{n-1}$, (a″) $\sum a_n$ is convergent, and (b′) $\liminf u_n < +\infty$. Using (a′) gives $u_p - (a_p + a_{p+1} + \cdots + a_{q-1}) \leqq u_q$ when $0 < p < q$. For each fixed p, taking inferior limits as $q \to \infty$ gives

$$u_p - (a_p + a_{p+1} + \cdots) \leqq \liminf_{q \to \infty} u_q$$

where the right member is either finite or $+\infty$. Taking superior limits as $p \to \infty$ gives

$$\limsup_{p \to \infty} u_p \leqq \liminf_{q \to \infty} u_q.$$

This implies that either $\lim u_p$ exists or $\lim u_p = +\infty$. The hypothesis (b′) eliminates the latter possibility and gives the conclusion that the sequence $u_1, u_2, \cdots$ is convergent.

Also solved by P. M. Anselone, Joshua Barlaz, Ranko Bojanic, G. U. Brauer, A. Césaro, Ward Cheney, T. Y. Chow, R. M. Conkling, R. J. Driscoll, Harley Flanders, L. C. Graue, D. S. Greenstein, C. A. Grimm, A. S. Hendler, Peter Henrici, S. S. Holland, Jr., J. B. Johnston, J. B. Kelly, J. H. B. Kemperman, W. A. Kistler, W. S. Lawton, A. E. Livingston, Marshall Luban, R. M. Meisel, E. J. Miller, D. J. Newman, E. N. Nilson, F. D. Parker, B. E. Rhoades, Michael Skalsky, D. P. Squier and David Zeitlin, Chih-yi Wang, L. E. Ward and L. E. Ward, Jr., and the proposer.

Necessary and Sufficient Condition for a Prime

4704 [1956, 584]. *Proposed by L. E. Clarke, University College of the Gold Coast*

If m is a positive integer greater than 1, show that a necessary and sufficient condition that m be prime is that, for every integer n,

$$\binom{n}{m} \equiv [n/m] \pmod{m}.$$

Solution by E. M. Wright, King's College, Aberdeen, Scotland. Given $|x| < 1$, we have

$$\sum_{n=m}^{\infty}\binom{n}{m}x^{n-m} = (1-x)^{-m-1},$$

$$\sum_{n=m}^{\infty}[n/m]x^{n-m} = (1+x+x^2+\cdots+x^{m-1})(1+2x^m+3x^{2m}+\cdots)$$

$$= (1-x)^{-1}(1-x^m)^{-1}.$$

Hence the condition of the problem is equivalent to the truth of the identical congruence

$$(1-x)^m \equiv 1-x^m \pmod{m}. \tag{1}$$

If m is a prime, this is well known (Hardy and Wright, *Theory of Numbers*, Theorems 75, 76) and almost trivial. If m is composite, let p be a prime divisor of m. Then (1) implies that $m \mid \binom{m}{p}$, and so $p \mid (m-1)(m-2)\cdots(m-p+1)$, which is impossible since $p \mid m$. Hence (1) is a necessary and sufficient condition that m be prime.

Also solved by D. M. Bloom, Leonard Carlitz, Leopold Flatto, A. S. Hendler, J. H. Hodges, J. H. B. Kemperman, D. C. B. Marsh, J. B. Muskat, C. D. Olds, Paul Schaefer, G. J. Smith, and the proposer.

Isomorphisms of the Additive Group

4705 [1956, 584]. *Proposed by Albert Wilansky, Lehigh University*

It is possible to construct an additive but discontinuous real function of one real variable. Is it possible to construct one with the property that $f(x)=f(y)$ implies $x=y$? (This would give an algebraic isomorphism of the additive group of the reals which is not continuous.)

Solution by G. E. Bredon, Harvard University. All possible additive functions are obtained by assigning arbitrary values to the elements of a Hamel basis over the rationals and extending the function by additivity to the reals. The isomorphisms of the additive group of the reals are obtained by mapping one Hamel basis in a one-to-one manner onto itself or some other basis, and all of these isomorphisms are discontinuous except the (trivial) ones of the form $x \to \alpha x$, α fixed.

Also solved by Ward Cheney, Harley Flanders, J. Horváth, H. H. Johnson, J. B. Johnston, J. B. Kelly, J. H. B. Kemperman, L. E. Ward, Jr., and the proposer.

Editorial Note. See Hamel's original discussion, *Eine Basis aller Zahlen und die unstetigen Lösungen der Funktionalgleichung* $f(x+y)=f(x)+f(y)$, Math. Ann. vol. 50, 1905, pp. 459–462.

Periodic Functions Having Incommensurable Periods

4706 [1956, 584]. *Proposed by S. W. Golomb, California Institute of Technology*

If $f(x)$ and $g(x)$ are real periodic functions of the real variable x, having incommensurable periods, is it possible for $f(x)+g(x)$ to be periodic?

I. *Solution by R. H. F. Denniston, University College, Ibadan, Nigeria.* Let

a, b, c be real numbers, no two being commensurable. For real x, let

$$d(x) = \begin{cases} 1 & \text{for } x = 0, \\ 0 & \text{for } x \neq 0; \end{cases}$$

$$f(x) = \sum_{m,n=-\infty}^{\infty} d(x - ma - nb) - \sum d(x - ma - nc);$$

$$g(x) = \sum d(x - mb - nc) - \sum d(x - ma - nb).$$

Then $f(x)$, $g(x)$, $f(x)+g(x)$ are periodic with periods a, b, c.

II. *Solution by Harley Flanders, University of California, Berkeley.* An affirmative answer is indicated by the following construction. Let a, b, c be real numbers which are linearly independent over the rationals and imbed these in a Hamel basis $\{a, b, c, \rho_i\}$ of the reals over the rationals. We define linear (over the rationals) real valued functions f and g by $f(a)=0$; $f(b)$, $f(c)$, $f(\rho_i)$ arbitrary; $g(b)=0$; $g(c)=-f(c)$; $g(a)$, $g(\rho_i)$ arbitrary. Then f has period a, g has period b, and $f+g$ has period c.

Also solved by G. E. Bredon, J. B. Kelly, J. H. B. Kemperman, Darald Rohb, and the proposer.

Magic Matrix

4707 [1956, 584]. *Proposed by Charles Fox, McGill University*

A magic matrix is one whose elements are the numbers of a magic square, *i.e.*, every row, column and diagonal has the same sum. (a) Show that a 3 by 3 magic matrix inverts into a magic matrix. (b) Can this result be extended to magic matrices of higher order?

Solution by F. D. Parker, Clarkson College of Technology, Potsdam, N. Y. Let A be a nonsingular magic matrix of any order, let $R(x)$ be a row matrix whose elements are all x, and $C(x)$ be a column matrix whose elements are all x, and let A' be the inverse of A. The condition that the rows of A all have the sum s is $AC(1)=C(s)$. Multiplying on the left by A' we have $C(1)=A'C(s)$, or the rows of A' have the same sum $1/s$. The corresponding analysis of the column sums goes from $R(1)A=R(s)$ to $R(1)=R(s)A'$, which shows that the columns of A' all have the same sum $1/s$.

It is not hard to show in a straightforward way that for $n=3$, if the diagonals of A also have the sum s, then the diagonals of A' have the same sum $1/s$. That this result does not necessarily extend to the fourth order is shown by the counter-example:

$$A = \begin{pmatrix} 2 & 3 & 1 & 2 \\ 2 & 1 & 5 & 0 \\ 3 & 0 & 2 & 3 \\ 1 & 4 & 0 & 3 \end{pmatrix}, \qquad A' = \frac{1}{64}\begin{pmatrix} 66 & -14 & 2 & -46 \\ 18 & 2 & -14 & 2 \\ -30 & 18 & 2 & 18 \\ -46 & 2 & 18 & 34 \end{pmatrix}.$$

Also solved by G. E. Bredon, D. C. B. Marsh, B. E. Mitchell, Chih-yi Wang, and the proposer.

Editorial Note. The fact that the rows of the inverse matrix have the same sum $1/s$ for all orders has been noted by A. Wilansky. See this MONTHLY, vol. 58, 1951, pp. 614–615 and vol. 63, 1956, pp. 652–653, where also the result is extended to certain infinite matrices.

Series Evaluation of Definite Integral

4708 [1956, 669]. *Proposed by H. A. Bender, University of Rhode Island*

Prove the following relationship:

$$\int_0^1 e^{c(1-x^2)}dx = \sum_{n=0}^{\infty} \frac{2^n c^n}{3\cdot 5\cdot\cdots\cdot(2n+1)}.$$

I. *Solution by Margaret M. LaSalle, Southwestern Louisiana Institute.* Repeated integration by parts gives:

$$\int_0^1 e^{c(1-x^2)}dx = 1 + \frac{2c}{3} + \frac{4c^2}{3\cdot 5} + \cdots + \frac{2^n c^n}{3\cdot 5\cdot\cdots\cdot(2n+1)} + \frac{2^{n+1}c^{n+1}}{3\cdot 5\cdot\cdots\cdot(2n+1)}\int_0^1 x^{2n+2}e^{c(1-x^2)}dx.$$

Since the integral on the right hand side is bounded, and the coefficient of the integral approaches zero as n approaches infinity, the result follows.

II. *Solution by L. A. Ringenberg, Eastern Illinois State College.* Expand the integrand in a series of powers of $(1-x^2)$, substitute $\sin\theta$ for x, and integrate termwise using Wallis' formula:

$$\int_0^1 e^{c(1-x^2)}dx = \int_0^1 \sum_{n=0}^{\infty} \frac{c^n(1-x^2)^n}{n!}\,dx = \int_0^{\pi/2} \sum_{n=0}^{\infty} \frac{c^n \cos^{2n+1}\theta d\theta}{n!}$$

$$= \sum_{n=0}^{\infty} \frac{c^n\cdot 2\cdot 4\cdot\cdots\cdot 2n}{n!\cdot 1\cdot 3\cdot\cdots\cdot(2n+1)} = \sum_{n=0}^{\infty} \frac{2^n c^n}{1\cdot 3\cdot\cdots\cdot(2n+1)}.$$

III. *Solution by D. C. B. Marsh, Colorado School of Mines.* Define

$$S(x) = \sum_{n=0}^{\infty} \frac{2^n c^n x^{2n+1}}{3\cdot 5\cdot\cdots\cdot(2n+1)}.$$

It is easily seen that $S(x)$ satisfies the linear differential equation

$$S'(x) - 2cxS(x) = 1.$$

Introduce the integrating factor e^{-cx^2} and integrate from $(x, S) = (0, 0)$ to $(1, S(1))$ to obtain the desired result in the form

$$e^{-c}S(1) = \int_0^1 e^{-cx^2}dx.$$

Also solved by D. M. Bloom, J. L. Boal and J. B. Muskat, Fred Brafman, Leonard Carlitz,

Tien Chi Chen, R. A. Costa, A. E. Danese, J. O. C. Ezeilo, G. B. Findley, Harley Flanders, J. S. Frame, Nathaniel Grossman, Emil Grosswald, R. R. Gutzman, A. B. Harper, Jr., Juris Hartmanis, Peter Henrici, Frank Herlihy, J. H. Hodges, A. F. Horne, A. R. Hyde, W. C. James, J. B. Johnston, Marshall Lubin, Y. L. Luke, R. T. Mahoney, R. M. Meisel, C. D. Olds, F. D. Parker, M. J. Pascual, Mary Payne, B. E. Rhoades, R. E. Shafer, Lawrence Shepp, F. C. Smith, J. R. Trollope, Chih-yi Wang, R. E. Williamson, and David Zeitlin.

Editorial Note. As pointed out by Lubin and Rhoades, above solutions I and III, virtually unchanged, yield

$$\int_0^1 x^r e^{c(1-x^a)}dx = \sum_{n=0}^{\infty} \frac{a^n c^n}{(a+r+1)(2a+r+1)\cdots(na+r+1)}, \quad r > -1, a > 0.$$

Isotomic Lines

4710 [1956, 669]. *Proposed by Hüseyin Demir, Zonguldak, Turkey*

Prove that if in a complete quadrangle inscribed in a circle (O) one pair of opposite sides are isotomic lines with respect to a triangle inscribed in (O), then the remaining pairs of opposite sides are also isotomic lines with respect to the same triangle.

I. *Solution by O. J. Ramler, Catholic University of America.* Using a system of conjugate coordinates we take the circle (O) as the unit circle, and on it points whose vector coordinates are T_1, T_2, T_3 as the inscribed triangle and t_1, t_2, t_3, t_4 as the inscribed complete quadrangle. Then the line t_2t_3 intersects the side T_2T_3 of the triangle in a point whose vector coordinate is

$$z = \frac{T_2T_3(t_2 + t_3) - (T_2 + T_3)t_2t_3}{T_2T_3 - t_2t_3}.$$

Similarly t_1t_4 meets side T_2T_3 where

$$z' = \frac{T_2T_3(t_4 + t_1) - (T_2 + T_3)t_1t_4}{T_2T_3 - t_1t_4}.$$

The hypothesis implies $T_3 - z = z' - T_2$ which becomes, upon making proper substitutions and simplifying

$$T_2^2T_3^2(T_2 + T_3 - s_1) + T_2T_3s_3 - (T_2 + T_3)s_4 = 0,$$

where s_1, s_3, s_4 are elementary symmetric functions of t_1, t_2, t_3, t_4. The symmetry of this result establishes the proposed theorem.

II. *Solution by Roland Deaux, Faculté Polytechnique, Mons, Belgium.* Circle (O) may be replaced by any conic Γ. Let ABC and $PQRS$ be the triangle and the quadrangle inscribed in Γ. By virtue of Desargues' theorem, the four conics Γ, (PQ, RS), (PR, QS), (PS, QR) determine on each side of ABC four pairs of an involution. This involution, defined by Γ and the isotomic lines PQ, RS, has for double points the midpoint and the point at infinity of the side. Hence the property.

Also solved by J. W. Clawson, N. A. Court, R. Goormaghtigh, Josef Langr, and the proposer.

RECENT PUBLICATIONS

EDITED BY RICHARD V. ANDREE, University of Oklahoma

All books for review should be sent directly to R. V. Andree, Department of Mathematics, University of Oklahoma, Norman, Oklahoma, and not to any of the other editors or officers of the Association.

Fundamental Concepts of Higher Algebra. By A. Adrian Albert. University of Chicago Press, 1956. ix+165 pp. $6.50.

The classic book of L. E. Dickson, *Linear Groups with an Exposition of the Galois Field Theory* (Teubner, 1901) has gained stature with the passing years. The treatment of linear groups was more complete than even the author probably realized, and the book has been, over the years, the major source of information on finite fields. It has inspired considerable research in the last half century, and furthermore, notations and points of view have changed. It is timely, therefore, to have an up-to-date exposition of the subject of finite fields using the slick proofs of modern algebra and embracing the latest results.

As the author states, the real reason for the existence of this book is the fifth and last chapter of 29 pages. The usual basic theorems on Galois fields are quickly developed. A proof is given of Gauss' theorem determining the integers m for which primitive roots modulo m exist. The Galois group of the finite field is determined. The familiar theorems of Serret and Dedekind on irreducible polynomials are given, and the concept of exponent to which an irreducible polynomial belongs is used to extend and clarify this theory. Now come many detailed results on the construction of irreducible polynomials. The book ends with a discussion of Dickson's method for computing primitive irreducible polynomials, and a list without proofs of eighteen known theorems.

In order to make the book essentially self-contained, the author has put in four preliminary chapters totaling 122 pages. These are entitled: I, Groups; II, Rings and Fields; III, Vector Spaces and Matrices; IV, Theory of Algebraic Extensions. In order to cover so much material in so short a space, the author has employed two devices. He has limited the theory to those topics which are necessary for his Chapter V, and he has in places drastically condensed the usual presentation. Thus Chapter III omits quadratic form theory, and all proofs of theorems on determinants.

It is thus evident that the title of the book is too comprehensive. In the reviewer's opinion the book is not well suited for use as a textbook with beginning graduate students. For those who have been over the material once, it would furnish an excellent review and summary of many parts of modern algebra. It is a scholarly piece of work, with some misprints, to be sure (but what book does not have some?), and is the last word on finite fields.

CYRUS COLTON MACDUFFEE
The University of Wisconsin

The Enjoyment of Mathematics. By Hans Rademacher and Otto Toeplitz. Princeton University, 1956, 204 pp. $4.50.

The first edition of *Von Zahlen und Figuren*, by Rademacher and Toeplitz, was published just twenty-seven years ago, but it is recognized today as a superlative piece of mathematical writing to which it is not inappropriate to attach the adjective "classical". Although the beauty of the German language may have escaped some English speaking people, the depth and subtle power of the mathematics were apparent to any serious reader.

The Enjoyment of Mathematics is a faithful translation into English of this classic to which the translator, Herbert Zuckerman, has added two fine chapters of his own. Because this is a translation of a well-known work we mention only a few of the topics discussed: The Sequence of Prime Numbers, Traversing Nets of Curves, Incommensurable Segments and Irrational Numbers, The Theory of Sets, Waring's Problem, Pythagorean Numbers and Fermat's Theorem, The Figure of Greatest Area with a Given Perimeter, Periodic Decimal Fractions, Curves of Constant Breadth. However a mere listing of the contents can give no hint of the simple but fundamentally deep character of the book.

Each chapter is a gem of mathematical exposition and the plan of attack is the same. The reader, who needs no more mathematical training than a good high school affords, is first introduced to some innocent-looking mathematical question. Next, with an absolute minimum amount of computation, he is carried deeper and deeper into the problem until the basic issues are made excitingly clear. Where final answers are still unknown (Fermat's Theorem, The Four-Color Problem, *etc.*), partial answers are obtained.

The Princeton University Press is to be congratulated for making this English edition available. It will not only stretch the imagination of the amateur, but it will also give pleasure to the sophisticated mathematician. It should prove extremely useful to teachers in quickening the mathematical spirit of students.

C. O. Oakley,
Haverford College

The Basic Concepts of Mathematics. Part I, *Algebra.* By Karl Menger, The Bookstore, Illinois Institute of Technology, 1957, v+91 pp.

In the words of the author, "*this book . . . is addressed to everyone desirous to clarify his ideas about some of the basic concepts and fundamental procedures of mathematics.*" The chapter headings indicate the procedures and concepts treated. They are: Numbers and Numerals, Facts and Formulas, Variables in General Statements, Unknowns in Equations, Variables in Description of Classes, Parameters in General Problems, General Expressions, Complex Numbers, Indeterminates in Polynomials.

A particular goal of the book is to resolve the many difficulties that arise through the use of the words "variable", "indeterminate", "parameter", "con-

stant". Professor Menger accomplishes this aim with charm and skill by means of numerous examples and also by references to the words' mathematical origin and subsequent history.

The book's tone is conversational and informal. Examples are discussed in detail and exercises are provided at appropriate places. It is a very readable introduction to mathematical semantics. The author has accomplished his objective in that this book should appeal greatly to the persons to whom it is addressed.

E. H. CRISLER
University of Notre Dame

Geometric Algebra. By E. Artin. Interscience, New York, 1957. x+214 pp. $6.00.

Most of this book is devoted to the study of algebraic structures arising from various geometries. The approach is algebraic rather than geometric. A good deal of the material was previously unavailable in English or in book form. The author has taken pains to write a wholly self-contained and readable exposition of this subject on the first year graduate level. He is to be congratulated on his success in this endeavor.

The author has adopted the wise course of including in Chapter I a large body of algebraic background material, some elementary and some not so elementary, which is used in the later portions of the book. The reader would do well to follow the author's suggestion to start with Chapter II and refer back to Chapter I when needed.

In Chapter II, affine geometry is introduced axiomatically and then coordinatized. Even here the approach is algebraic, *i.e.* a Desarguian plane is obtained by assuming that there exist "enough" translations and dilatations. The Pappus and Desargues theorems are then considered. The connection between ordered geometries and ordered coordinate fields is studied using material from Chapter I on ordered fields. The fourth harmonic point construction is given as a computation in the coordinate field. This computation gives the criterion for the field to have characteristic 2 and is also used, together with a theorem of Hua (proved in Chapter I), to prove a generalization of von Staudt's theorem concerning maps of a line preserving harmonic points.

Projective geometry is introduced, already coordinatized, as the geometry of subspaces of a vector space over a field. The author then proves the fundamental theorem connecting the collineations of the geometry with semilinear transformations of the underlying vector space. This, in turn, is used to obtain an exact diagram which concisely shows the connections between the various groups involved (collineation group, general linear group, field automorphism group, *etc.*). Chapter II closes by relating the affine and projective geometries.

Symplectic and orthogonal geometries are introduced in Chapter III by considering a vector space (over a commutative field) with an inner product. There is a nice concise section in Chapter I concerning pairings and duality which is necessary for most of the results of Chapter III. Chapters II and III are independent.

Orthogonal and symplectic geometries are studied together, until the former is identified as an orthogonal sum of lines and the latter (in the nonsingular case) as a sum of hyperbolic planes. Looking at these geometries separately, he considers a number of results on how the group of isometries acts on the geometry, *e.g.* an isometry on orthogonal geometry is almost determined by its action on a hyperplane, the symplectic group contains only rotations, *etc.* Chapter III ends with the computation of the orders of the various groups in case the field is finite.

Chapter IV, The General Linear Group, begins with a discussion of determinants over noncommutative fields. This is used to study the relations between the general linear group, the group generated by transvections, their centers, centralizers, *etc.* The orders of these groups are computed in the case of a finite field.

A similar study is undertaken in Chapter V on the structures of the symplectic and orthogonal groups. The theory here is not as yet complete, so the results are necessarily more complicated. It is necessary to study a number of special cases.

All in all, the book is delightful. The author's choice of material is sound; for example, he avoids the peculiarities of non-Desarguian planes, orthogonal geometry over a field of characteristic 2. His presentation is clear and well motivated. Indeed, one of the best things about the book is that it sounds like the author talking, complete with chatty remarks (see his comment on matrix methods in Chapter I). This is a book which can and will be enjoyed by a large section of the mathematical public.

JAMES P. JANS
Ohio State University

Elements of Mathematics (*2nd Edition*). By Helen Murray Roberts and Doris Skillman Stockton. Addison-Wesley, Reading, Mass. 1956, x+308 pp. $3.50.

The book, according to the preface, was written for the purpose of preparing students deficient in secondary school mathematics for college mathematics. It fulfills this purpose rather well on the whole.

There is a clear and careful development of basic ideas which are too often assumed to be a part of a student's background. Examples of this are plentiful; the discussion of our number system, the necessary extensions of the system from positive integers, to negative integers, to rational numbers, *etc.*, the discussion of the decimal number system are all well organized and meaningful to the student. There is a thorough and clear explanation of dependent and inconsistent equations. The confusion concerning rational exponents is handled very well.

This reviewer would question the use of the word "remainder" as a substitute for "difference" in subtraction. It does not make the process more clear and could cause confusion with other uses of the word remainder. Also the practice of avoiding the use of the standard names for Associative, Commutative, and

Distributive Laws does not seem logical, since they are properly descriptive terms.

The text defines a monomial in the following manner: "Any combination of numbers and literal numbers formed only by multiplication and division has been called a term. Frequently a single term is called a monomial." Then it defines a polynomial as "the sum of two or more monomials." This meaning is usually reserved for the word *multinomial*, with *polynomial* being defined in a much more restrictive sense.

On the whole, however, the care used in the development of fundamental concepts places the book among the best recent efforts.

DONALD R. LINTVEDT
Upsala College

Irrational Numbers. By Ivan Niven. Carus Mathematical Monograph No. 11. The Mathematical Association of America (distributed by Wiley, New York), 1956. 164 pp. $3.00.

The investigation of the nature of irrational numbers may well serve as a guide to number theory. After all, a good deal of the theory of divisors and prime numbers in Euclid's *Elements* may have been obtained with the aim of understanding the irrationality of surds. Nowadays, since we distinguish within the domain of irrational numbers between algebraic and transcendental ones, a discussion of irrational numbers must necessarily include a fair amount of the theory of algebraic number fields. A book on irrational numbers has therefore a natural place in a collection of expository articles such as the *Carus Monographs*.

The present book fulfills its purpose excellently. It covers the most important facts of the theory itself and contains references and suggestions for further reading concerning those topics, which surpass in extent and intricacy the scope of such a small volume. At the end of each of the ten chapters the author has appended notes of historical and bibliographical character and informative hints about related and more complicated matters. Commendable also is the list of notations and the glossary with skillfully worded explanations at the end of the book.

The contents of the book, chapter by chapter are: I. Rationals and Irrationals. The noncountability of the irrationals is shown by an argument of measure. Decimal expansions and the Cantor expansion give criteria for irrationality. II. Irrationality of $\cos r$ for rational $r \neq 0$ is proved by a variant of the author's well-known proof of the irrationality of π, using Hermite's device $F(x) = \sum(-1)^{\nu} f^{(2\nu)}(x)$, where $f(x)$ is a suitable polynomial. The proof that e is transcendental is also given here since all preparations are at hand. Chapter III deals with the algebraic nature of the roots of unity and with a result of D. H. Lehmer about $\cos(2\pi k/n)$. IV. The pigeon-hole principle is used for the discussion of the approximation of irrationals by rationals and also simultaneous approximation of several irrationals. Chapter V gives a concise little exposition

of continued fractions, especially periodic ones. Chapter VI discusses the goodness of approximation with Hurwitz's theorem about best approximation. It also contains two proofs of the uniform distribution modulo 1 of the sequence $\{n\xi\}$, where ξ is irrational. Chapter VII on algebraic and transcendental numbers contains, in particular, Liouville's theorem. The Thue-Siegel theorem is also mentioned in the notes. Chapter VIII deals with normal numbers. Borel's definition is compared with the more natural one which considers the frequency of blocks of k digits. That both definitions are equivalent is shown by theorems of Pillai and of Niven and Zuckerman. The last two chapters are less elementary. Chapter IX gives the general Lindemann theorem. For the algebraic tools needed here the book refers to the *Carus Monograph* by Pollard. In Chapter X, Gelfond's and Schneider's theorem that α^β is transcendental for algebraic α and β (with the obvious trivial exceptions) is proved. The proof follows Siegel's presentation of Gelfond's proof, but gives many more details than Siegel's concise version. The notes to this last chapter refer also to other classes of transcendental numbers which are considered, in particular, in Siegel's treatise on transcendental numbers.

This survey shows that all important features of the theory of irrational numbers are treated or at least touched upon. Niven's book is written in a clear and attractive style. It is a very valuable addition to the still too small number of expository books in our science.

HANS RADEMACHER
University of Pennsylvania

BRIEF MENTION

Theory of Lie Groups. By Claude Chevalley. Princeton, 1957. viii+217 pp. $2.75.

It is welcome news that this out-of-print volume is again available in the paper back edition.

Integration. By Edward J. McShane. Princeton, 1957. viii+394 pp. $2.95.

Another welcome reissue which needs no further review other than to announce its availability (paper back).

Applied Mathematics in Chemical Engineering. By H. S. Mickley, T. K. Sherwood, C. E. Reed. McGraw-Hill, 1957. xii+413 pp. $9.00.

This reviewer is rather disappointed that so few *modern* additions appear in the revision of this excellent test. No mention, for example, is made of the modern applications of matric algebra to chemical engineering. On the other hand, the inclusion of the Laplace transform and of finite differences, and some work on numerical methods is most welcome.

Descriptive Geometry. By Steve M. Slaby. Barnes and Noble, 1956, 353 pp. $2.25.

Automation: Its Purpose and Future. By M. Pyke. Philosophical Library, 1956, 191 pp. $10.00.

The Electrical Production of Music. By Alan Douglas. Philosophical Library, 1957, 223 pp. $12.00.

NEWS AND NOTICES

EDITED BY EDITH R. SCHNECKENBURGER, University of Buffalo

Readers are invited to contribute to the general interest of this department by sending news items to Edith R. Schneckenburger, University of Buffalo, Buffalo 14, New York. Items must be submitted at least two months before publication can take place.

ASEE MATHEMATICS DIVISION

The Mathematics Division of the American Society for Engineering Education met on June 17–21, 1957 at Cornell University. Three well-attended sessions were held, including a joint conference with the Upper New York State Section of the Mathematical Association of America. The following new officers of the Division were elected at the annual business meeting: Chairman, Professor C. R. Wylie, Jr., University of Utah; Vice-Chairman, Professor H. A. Giddings, New York University; Secretary, Professor W. E. Restemeyer, University of Cincinnati; Directors, Professors W. G. Warnock, Rensselaer Polytechnic Institute, G. B. Thomas, Massachusetts Institute of Technology, C. O. Oakley, Haverford College; Representative to ASEE General Council, Professor Haim Reingold, Illinois Institute of Technology. The next annual meeting of the Mathematics Division will be held in June 1958 at Berkeley, California. For further information write to Professor W. E. Restemeyer, University of Cincinnati.

NATIONAL SCIENCE FOUNDATION PUBLICATIONS

National Science Foundation publications are available to all scientists who have need for them. Lists of publications are available from the Foundation. Requests should be addressed to the Publications Office, National Science Foundation, Washington 25, D. C.

PRELIMINARY ACTUARIAL EXAMINATIONS PRIZE AWARDS

The winners of the prize awards offered by the Society of Actuaries to the nine undergraduates ranking highest on the score of Part 2 of the 1957 Preliminary Actuarial Examination are as follows:

First Prize of $200.

Solovay, Robert	Harvard University

Additional Prizes of $100 each.

Flittie, John H.	Drake University
Gardner, John R.	University of Toronto
Kandall, Geoffrey A.	Princeton University
Lakser, Harry	University of Manitoba
Lichtenbaum, Stephen	Harvard University
Posner, Paul	Princeton University

Sadowsky, George	Harvard University
Zvengrowski, Peter D.	Rensselaer Polytechnic Institute

The Society of Actuaries has authorized a similar set of nine prizes for the 1958 examinations on Part 2.

The Preliminary Actuarial Examinations consist of the following three examinations:

Part 1. Language Aptitude Examination.
(Reading comprehension, meaning of words and word relationships, antonyms, and verbal reasoning.)

Part 2. General Mathematics Examination.
(Algebra, trigonometry, coordinate geometry, differential and integral calculus.)

Part 3. Special Mathematics Examination.
(Finite differences, probability and statistics.)

The 1958 Preliminary Actuarial Examinations will be prepared by the Educational Testing Service under the direction of a committee of actuaries and mathematicians and will be administered by the Society of Actuaries at centers throughout the United States and Canada on May 14, 1958. The closing date for applications is April 1, 1958.

PERSONAL ITEMS

Dr. M. H. Martin, Executive Secretary, Division of Mathematics, National Academy of Sciences-National Research Council, has resigned. Professor H. W. Kuhn, Bryn Mawr College, is replacing Dr. Martin.

Professor Everett Pitcher, Lehigh University, received the honorary degree of Doctor of Science from Western Reserve University in June 1957 on the twenty-fifth anniversary of his degree of Bachelor of Arts.

Antioch College: Professor Parker Hamilton has retired as Head of the Department of Mathematics; Associate Professor J. H. Blau has been appointed Chairman of the Department; Associate Professor W. K. Smith, Bucknell University, has been appointed Associate Professor; Assistant Professor Daniel Sokolowsky has been promoted to Associate Professor.

Brown University: Assistant Professors David Gale and John Wermer have been promoted to associate professorships; Dr. W. D. Barcus has been promoted to Assistant Professor; Mr. D. K. Harrison, Intern in Mathematics, has been appointed Instructor; Mr. A. M. Duguid, Cambridge University, has been appointed Instructor for the academic year 1957–58; Dr. E. H. Brown, Jr., University of Chicago, is holding a research associateship during the academic year 1957–58 under the support of ONR; Professor Herbert Federer has received a grant from the Sloan Foundation and is on leave of absence during 1957–58; Professor David Gale is on leave of absence with the RAND Corporation, Santa Monica, California.

Cornell University: Visiting Assistant Professor S. S. Abhyankar, Columbia University, has been appointed Assistant Professor; Associate Professor I. N. Herstein, University of Pennsylvania, has been appointed Associate Professor; Dr. Alfred Aeppli, Assistant, Swiss School of Technology, has been appointed Instructor; Dr. G. R. Livesay, Research Associate at the University, has been appointed Visiting Assistant Professor; Associate Professor Daniel Gorenstein, Clark University, has been appointed Visiting Lecturer; Assistant Professor H. D. Block has been appointed Associate Professor, Department of Mechanics; Dr. Walter Feit has been promoted to Assistant Professor; Associate Professors G. A. Hunt and Paul Olum have been promoted to Professors; Mr. Hung Ching Chow, Acting Director, Institute of Mathematics, Academia

Sinica, Taipei, Taiwan, China, is a Visiting Scholar on a Mundt, Smith, and Fulbright grant; Assistant Professor C. S. Herz is on leave during 1957–58 at the Institute for Advanced Study. An Institute for Symbolic Logic, sponsored by the American Mathematical Society under a grant from the National Science Foundation, was held at the University during the summer of 1957.

Illinois Institute of Technology: The Carnegie Corporation of New York has awarded the Institute a three years grant which will be used by Professor Karl Menger in developing a new approach to teaching of mathematics.

Montana State College: Associate Professor A. L. Hess has been promoted to Professor; Assistant Professor J. E. Whitesitt has been promoted to Associate Professor.

New York University, School of Commerce, Accounts, and Finance: Associate Professor H. E. Wahlert has been promoted to Professor; Mr. Stefan Bauer-Mengelberg has accepted a fellowship in music from the Frank Huntington Beebe Fund; Mr. Nathan Newman, Cooper Union, has been appointed Instructor.

University of Nebraska: Assistant Professor W. E. Mientka, University of Nevada, and Mr. Hubert Schneider, University of Munster, have been appointed Assistant Professors; Mr. S. E. Bohn and Mr. Konrad Suprunowicz, Graduate Assistants at the University, have been appointed Instructors; Dr. G. C. Cree, Dr. M. L. Keedy, and Dr. D. W. Miller have been promoted to Assistant Professors.

Professor C. B. Allendoerfer, who is on leave of absence from the University of Washington for the year 1957–58, is a Fulbright lecturer at Cambridge.

Dr. K. A. Brons, University of Illinois, has accepted a position as applied science representative for International Business Machines Corporation, River Forest, Illinois.

Mr. P. L. Chessin, Senior Engineer, Westinghouse Electric Corporation, Baltimore, Maryland, has received a National Science Foundation Science Faculty Fellowship and is at the Institute for Fluid Dynamics and Applied Mathematics, University of Maryland.

Associate Professor N. E. Dodson, Lenoir Rhyne College, has been appointed Assistant Professor at Wittenberg College.

Assistant Professor A. V. Fend, New Mexico College of Agriculture and Mechanic Arts, is now employed in Operations Research, Technical Operations, Monterey, California.

Mr. D. A. Franks, Chief, Ordnance Guided Missiles School, Internal Guidance Section, SAM Division, Redstone Arsenal, Huntsville, Alabama, has a position as a mathematician at Westinghouse Electric Corporation, Air Arm Division, Baltimore, Maryland.

Dr. A. J. Goldstein, Polytechnic Institute of Brooklyn, has accepted a position as a member of the technical staff, Bell Telephone Laboratories, Murray Hill, New Jersey.

Dr. Donald Greenspan, Research Engineer, Systems Development Laboratories, Hughes Aircraft Company, Culver City, California, has been appointed Assistant Professor at Purdue University.

Dr. Juris Hartmanis, Cornell University, has been appointed Assistant Professor at Ohio State University.

Dr. J. B. Johnston, Cornell University, has been appointed Assistant Professor at the University of Kansas City.

Dr. R. F. King, Argonne National Laboratories, has accepted a position as Scientist, Computer Division, Midwestern Universities Research Association, Madison, Wisconsin.

Mr. H. G. Loomis, Pennsylvania State University, has been appointed Instructor at Amherst College.

Dr. Joseph Mayer, Department of Economics, Miami University, has been appointed Professor of Mathematics at the University.

Dr. W. O. J. Moser, University of Saskatchewan, has been promoted to Assistant Professor.

Dr. J. P. Nash, University of Illinois, has accepted a position as Manager, Information Processing Division, Missile Systems Division, Lockheed Aircraft Corporation, Sunnyvale, California.

Dr. J. W. Neuberger, University of Texas, has been appointed Instructor at Illinois Institute of Technology.

Dr. J. C. C. Nitsche, Technical University of Berlin, has been serving as Visiting Associate Professor at the Institute of Technology, University of Minnesota, since February, 1957.

Dr. L. B. Rall, Mathematician, Shell Development Company, Emeryville, California, has been appointed Associate Professor at Lamar State College of Technology.

Professor Edward Saibel, Department of Mechanics, Carnegie Institute of Technology, has been appointed Visiting Professor of Mechanics at Rensselaer Polytechnic Institute.

Assistant Professor R. L. San Soucie, University of Oregon, has accepted a position as Head, Mathematics Group, Applied Research Department, Sylvania Electric Products, Buffalo, New York.

Dr. E. J. Schweppe, University of Nebraska, has been appointed Assistant Professor at Iowa State College.

Mr. Franklin Sheehan, San Francisco State College, has a position as Operations Analyst with Technical Operations, Monterey, California.

Dr. N. B. Stein, National Science Foundation Fellow, Cornell University, has been appointed Instructor at Yale University.

Dr. R. D. S. Tuan, Head, Department of Mathematics and Physics, Tougaloo Southern Christian College, has been appointed Professor at Indiana Technical College.

Dr. W. F. Whitmore is now on leave of absence from the Operations Evaluation Group to serve as Chief Scientist, Special Projects Office, Bureau of Ordnance, Department of the Navy, Washington, D. C.

Mr. J. W. Young, Jr., National Security Agency, has accepted a position as Chief, Advanced Systems Research Section, Electronics Division, National Cash Register Company, Hawthorne, California.

Mr. J. R. Ziegler, University of California at Los Angeles, has a position as Manager, Hawthorne Branch Electronic Sales, National Cash Register Company, Hawthorne, California.

Rev. L. H. Dubé, Ottawa University, died on May 29, 1957. He was a member of the Association for thirty-seven years.

Dean J. A. Hardin, Centenary College, died in August, 1955. He was a member of the Association for thirty-three years.

Mr. J. V. Longenecker, Wichita, Kansas, died on March 15, 1956. He was a member of the Association for thirty-one years.

Assistant Professor D. S. Nathan, City College of the City of New York, died on July 1, 1957.

Dr. A. R. Schweitzer, Lake Forest, Illinois, died on June 12, 1957. He was a charter member of the Association.

Professor R. L. Westhafer, New Mexico College of Agriculture and Mechanic Arts, died on July 3, 1957. He was a member of the Association for twenty-one years and served as an officer of the Southwestern Section for six years.

Professor Emeritus A. E. Whitford, Alfred University, died on April 14, 1957. He

was a charter member of the Association.

Dr. G. H. Wilson, Furlong, Pennsylvania, died on December 4, 1956. He was a member of the Association for twenty-three years.

THE MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

NEW MEMBERS

Professor H. M. Gehman, Secretary-Treasurer, announces that the following 103 persons have been elected to membership by the Board of Governors on applications duly certified.

WILLIAM R. ALFORD, Student, The Citadel.

RICHARD A. BACH, Student, Harpur College.

WILLIAM BENDER, Ph.D. (Yale) Professor & Head, Department of Physics, University of South Dakota.

MICHAEL BERNKOPF, B.A. (Dartmouth) Grad. Student, Columbia University.

MICHAEL H. BERNSTEIN, Student, Princeton University.

MRS. LAVERNA BOYLES, M.Ed. (Oklahoma) Teacher, Taft Junior High School, Oklahoma City, Okla.

EMIL W. BROWN, Traffic Engr., Mountain States Tel. & Tel. Co., Phoenix, Ariz.

THOMAS E. CALDWELL, B.A. (Illinois) Teaching Asst., University of Arizona.

FRANCIS P. CALLAHAN, JR., M.A. (Columbia) Math., Philco Corp., Philadelphia, Pa.

TIEN C. CHEN, Ph.D. (Duke) Asso. Physicist, I.B.M. Research Lab., Poughkeepsie, N. Y.

VIRGINIA B. CHRISTIAN, M.S. in Ed. (Eastern Illinois) Instr., Eastern Illinois State College.

DONALD A. COLE, Student, Georgia Institute of Technology.

GEORGE COPP, Ph.D. (Texas) Asso. Professor, North Texas State College.

RICHARD M. CORRY, Field Engr., Ebasco Services, Estacada, Ore.

GLENN R. COWGILL, B.S. (Kent) Math., National Advisory Committee for Aeronautics, Cleveland, Ohio.

GEOFFREY CROFTS, B.Com. (Manitoba) Asso. Professor of Actuarial Science, Occidental College.

JOHN H. DAUWALDER, Student, Occidental College.

JOHN E. DENES, B.S. (C.C.N.Y.) Statistician, National Industrial Conference Board, New York, N. Y.

DOMER V. DOUGHERTY, M.Ed. (Oklahoma) Asso. Professor, Phillips University.

FRANCIS A. DUNN, Student, LaSalle College.

EARL O. EMBREE, M.S. (Illinois) Instr., Morgan State College.

GEORGE EPSTEIN, M.S. (Illinois) Res. Asst., University of California at Los Angeles.

HARRY D. EYLAR, B.A. (Montana S.U.) Grad. Asst., Montana State University.

ANN FAUCHALD, Student, University of Washington.

RICHARD W. FELDMANN, JR., Student, University of Buffalo.

JOHN R. FLORENCE, JR., B.S. in M.E. (Colorado) Instr., University of Colorado.

MARY E. GABBERT, Student, Baylor University.

AUDLEY D. GASTON, JR., B.A. (Texas) Austin, Texas.

HAL M. GILMORE, M.A. (Western Kentucky) Instr., Illinois State Normal University.

JUDITH GLEASON, Student, Baylor University.

FRED G. GUSTAVSON, Student, Rensselaer Polytechnic Institute.

KATHRYN B. HANCHON, M.S. (Michigan) Engg. Asst., General Electric Co., Cincinnati, O.

PATRICIA A. HANDRICKEN, Student, Regis College.

HAROLD B. HANES, JR., Student, Texas Christian University.

DUNSTAN HAYDEN, O.S.B., Ph.B. (Catholic) Teacher, St. Anselm's Priory School, Washington, D. C.

JAMES B. HILDEBRAND, Student, Albion College.

SHELBY K. HILDEBRAND, B.A. (North Texas) Teaching Fellow, North Texas State College.

JOEL W. HOLLENBERG, Student, Cooper Union.

HOWARD T. HUMPHREY, Student, University of Buffalo.

RORA F. IACOBACCI, M.A. (Fordham) Instr., Manhattanville College of the Sacred Heart.

MRS. SONDRA O. JAFFE, Student, Brooklyn College.

JAMES H. JORDAN, B.S. (Southern Oregon) Grad. Asst., University of Oregon.

ALLEN J. KAPLAN, Student, Boston University.

STANLEY KAPLAN, Student, Cornell University.

ARTHUR F. KAUPE, JR., Student, Carnegie Institute of Technology.

REV. JOHN F. KELLER, S.J., M.S. (St. Louis) Asst. Professor & Chm., Department of Mathematics, Loyola University, La.

LOUIS J. KIJEWSKI, Student, LaSalle College.

GEORGE D. KING, M.A. (Alabama) Hd., Department of Mathematics, Brevard College.

JACOB O. KOEHL, Student, Agricultural and Mechanical College of Texas.

ELAINE H. KOPPELMAN, Student, Brooklyn College.

KENNETH S. KRETSCHMER, M.S. (Stanford) Grad. Student, Carnegie Institute of Technology.

DON R. LICK, B.S. (Michigan S.U.) Grad. Asst., Michigan State University.

JOE LIPMAN, Student, University of Toronto.

EDMONDO M. MASTROIANNI, M.S. (St. John's U.) Math., I.B.M. Corp., Endicott, N. Y.

ROBERT L. MCFARLAND, Student, Case Institute of Technology.

EARL H. MCKINNEY, M.S. (Pittsburgh) Instr., University of Pittsburgh.

BENJAMIN H. MCLEMORE, JR., M.A. (Illinois) Asso. Professor, Jackson State College.

RICHARD J. MILLER, M.A. (Syracuse) Instr., Clarkson College of Technology.

ENNIS J. MONTELLA, M.A. (Boston C.) Asst. Professor, Merrimack College.

INA MORAN, Student, Seton Hill College.

MERLYND K. NESTELL, Student, Emmanuel Missionary College.

DEWAYNE S. NYMANN, Student, Iowa State Teachers College.

ENUENWEMBA OBI, B.S. (Kansas City) Grad. Asst., University of Nebraska.

GARY H. OSBORN, Student, State College of Washington.

MRS. PATRICIA M. OVERDEER, M.S. (Delaware) Instr., Pennsylvania State University.

CLARENCE W. PATTY, B.S. (Georgia) Grad. Student, University of Georgia.

LYMAN C. PECK, Ph.D. (Ohio State) Asso. Professor, Ohio Wesleyan University.

WILLIAM F. POHL, Student, University of Chicago.

MATTHIAS F. REESE, B.S. (Houston) Houston, Texas.

WAYNE A. RHEA, Student, Baylor University.

HUBERT L. RICHARDS, Student, Eastern Kentucky State College.

MARY E. RIETMAN, Student, West Texas State College.

MRS. ETHEL A. ROBINSON, M.A. (Stanford) Instr., Fresno State College.

ERNEST L. ROETMAN, Student, University of Minnesota.

FRED D. ROSE, Student, University of Georgia.

REUBEN I. SANDLER, Student, Reed College.

ROBERT L. SAN SOUCIE, Ph.D. (Wisconsin) Asst. Professor, University of Oregon.

RICHARD E. SARBER, B.S.M.E. (Indiana Tech. C.) Instr., Indiana Technical College.

RONALD A. SCHAUFELE, B.Ed. (Alberta) Grad. Student, University of Washington.

RAYMOND H. SCHULZ, B.S. (Sam Houston S.C.) Grad. Student, Sam Houston State College.

MACEO T. SCOTT, M.A. (Columbia) Math., Flight Simulation Lab., White Sands Proving Ground, N. Mex.

M. THERESE SHEEHAN, A.B. (Dunbarton) Math., Army Map Service, Washington, D. C.

OSCAR D. SHELLEY, Res. Engr., North American Aviation, Downey, Calif.

ARNOLD SINGER, Student, Yeshiva University.

MARVIN B. SLEDD, Ph.D. (M.I.T.) Professor, Georgia Institute of Technology.

DONALD L. SMITH, Ph.D. (Rochester) Quality Engr., Eastman Kodak Co., Rochester, N. Y.

J. DAVID STANBERRY, Minneapolis, Minn.

PAULINE P. STEINBECK, Student, Central College, Mo.

JACK H. TAUB, Student, Rutgers University.

ROBERT G. TOBEY, Student, College of Wooster.

MARTIN L. USSERY, Student, Arkansas Polytechnic College.

SHERMAN B. VANAMAN, JR., M.S. (Kentucky) Chm., Department of Mathematics, Carson-Newman College.

ELDON J. VOUGHT, Student, Manchester College.

M. CURTIS WALKER, M.A. (Michigan) Math., General Motors Institute.

WILLIAM R. WHEELER, Student, Gannon College.

WILLIAM H. WHITEHEAD, III, Student, Ursinus College.

JACOB A. WILLIAMS, M.A. (Denver) Res. Math., Denver Research Institute.

ROBERT J. WINTERBOTTOM, III, Student, Ursinus College.

ALFRED H. WITTE, JR., Student, University of Nebraska.

ELAINE G. YODICE, B.A. (Hunter) Teaching Asst., University of Wisconsin.

JAMES M. YOHE, Student, DePauw University.

ROBERT R. A. YOUNG, Student, University of New Mexico.

HORAN B. ZARIAN, M.A. (California) Topographic Computer, U. S. Army.

THE APRIL MEETING OF THE METROPOLITAN NEW YORK SECTION

The sixteenth annual meeting of the Metropolitan New York Section of the Mathematical Association of America was held at Hunter College, New York City, April 27, 1957. Dean Mina Rees of Hunter College, Chairman of the Section, opened the meeting and presided at the morning session which was devoted to papers on some mathematical developments related to computer applications. Dr. Irving Dodes, High School Vice-Chairman of the Section, was moderator for the afternoon session which was devoted to papers on curriculum trends in high school and college mathematics. There were 122 persons in attendance, including 94 members of the Association.

The following officers were elected for the years 1957–59: Chairman, Professor J. N. Eastham, Cooper Union; Vice-Chairman for Colleges, Professor E. R. Stabler, Hofstra College; Vice-Chairman for High Schools, Dr. Reinhold Walter, Manhattan School of Aviation Trades; Secretary, Dr. Azelle B. Waltcher, Hofstra College; Treasurer, Mr. Aaron Shapiro, Midwood High School, Brooklyn.

At the business meeting reports were given by the Treasurer, the Committee on Contests and Awards, and the Committee on Coordination of Mathematics Training. The motion was unanimously passed that the Treasurer of the Section, Mr. Aaron Shapiro, receive an expression of appreciation and gratitude for his service to the Section.

At the Executive Committee Meeting of the Metropolitan New York Section in November, 1956 a resolution was adopted concerning the willingness of the Committee on Contests and Awards to work with the National Contest Committee. As a result of this action, the Committee on Contests and Awards brought to the annual meeting the following resolution, which was unanimously adopted:

Whereas, the Mathematical Association of America has set up a National Committee on Contests, and

Whereas, the Committee on Contests and Awards of the Metropolitan New York Section has been conducting annual contests since 1949, and

Whereas, the contest booklets and experience of the Metropolitan New York Section have been used, and are being used, by the Contest units in eight or nine states and provinces,

Therefore, be it resolved that the Metropolitan New York Section offers the services of its Committee on Contests and Awards to the National Committee in the promotion of the objectives of the annual high school contest.

The By-Laws of the Section were amended by unanimous vote to provide for two-year terms of office for the Section officers.

The Committee on Coordination of Mathematics Training of the Metropolitan New York Section formulated three resolutions which received the approval of the Executive Committee at the November 1956 meeting. The resolutions, which follow, were adopted by the Section at the annual meeting.

1. The Metropolitan New York Section of the M.A.A. recommends to all high school students who plan to study calculus in college that they should have at least three years of high school mathematics, and the Section recommends to all high school advisors that they strongly advise this course of study.
2. The Metropolitan New York Section of the M.A.A. recommends to the colleges that three years of high school mathematics be required as a minimum for all college entrants intending to study the calculus, and that high school mathematics through advanced algebra be required as a minimum for all college entrants planning courses of study in mathematics, science, or engineering.
3. The Metropolitan New York Section of the M.A.A. recommends to the high schools that prospective college entrants be separated from other students in high school mathematics classes wherever feasible, and that the respective teaching methods and course contents be chosen in accordance with the objectives for each group of students.

It was moved and unanimously carried that the Metropolitan New York Section go on record as approving the sentiments expressed in a telegram sent by Dean Mina Rees to Chancellor John P. Myers, New York State Board of Regents. The telegram stated in effect the opinion that an increase in the course requirements in professional education courses without consideration of the adequacy of subject matter preparation would weaken the preparation of teachers, and would influence adversely serious efforts of educational leaders to improve the quality of instruction. The opinion was also expressed in this telegram, that faculty leaders outside of professional education circles should be called upon to participate in decisions which affected the course requirements in the preparation of teachers.

The following papers were presented at the meeting:

1. *The UNIVAC election forecasts*, by Dr. M. A. Woodbury, College of Engineering, New York University.

The mathematical and statistical problems involved in applying probability models to forecasting the 1952, 1954, and 1956 elections using an electronic computer are discussed. The problems of data selection and rejection, transformation into appropriate form, inferences, and decisions based on the probability model are set forth.

2. *Linear programming*, by Dr. Philip Wolfe, Princeton University.

An example of a production scheduling problem is taken as typical of the management problems formulable in the linear programming model. Procedures for converting such problems into computationally accessible problems are illustrated by means of the example. The theory of positive linear dependence is sketched and applied to the linear programming problem, and used to motivate the computational algorithm of the simplex method.

3. *New trends in the undergraduate college mathematics curriculum*, by Dr. Seymour Schuster, Polytechnic Institute of Brooklyn.

The traditional college curriculum in mathematics is briefly analyzed with regard to its present shortcomings. Particular emphasis is given to the failure of the traditional program to serve the principles of liberal education, which require fluency in language and mathematics. The apparent trends which seek revision of the curriculum are described through a consideration of the efforts of the M.A.A. Committee on the Undergraduate Mathematical Program and the many faculties which are experimenting with their respective curricula. The influence of several of the modern texts is discussed through a brief analysis of each. Finally, hope is extended for a trend which would

manifest itself in a concerted effort (with all departments joining forces) to assume responsibilities to the discipline of language, as well as mathematics—since communication depends on both.

4. *Preparing the bright student in twelfth year high school mathematics*, by Mr. H. D. Ruderman, The Bronx High School of Science.

Students preparing for the Advanced Standing Examination in Mathematics must take analytic geometry and calculus. For those not taking the Examination, the following five criteria might be considered as a basis for selection of the course. The mathematics for the course should: 1) emphasize ideas having wide and important applications, 2) contain a good sampling of significant mathematical ideas, 3) aim at a better understanding of the foundations of mathematics, 4) be learnable without undue hardship but rather with considerable interest and excitement on the part of the student, 5) be teachable by the best prepared teachers in every high school, again without undue hardship but rather with interest and enthusiasm. At The Bronx High School of Science there is a course that meets these conditions. The first half consists of advanced algebra with a large portion of differential calculus limited to algebraic functions. The second half follows closely chapters 1 to 6 and 12 of "Principles of Mathematics" by Allendoerfer and Oakley published by McGraw-Hill.

AZELLE B. WALTCHER, *Secretary*

THE APRIL MEETING OF THE MISSOURI SECTION

The annual meeting of the Missouri Section of the Mathematical Association of America was held on April 27, 1957, at Southeast Missouri State College, Cape Girardeau, in conjunction with a meeting of the Missouri Council of Teachers of Mathematics. Professor W. R. Utz, Vice-Chairman of the Section, presided at the morning session of the Section. Professor R. J. Michel, Chairman of the Section, presided at the business meeting and the joint afternoon session. There were 54 members in attendance.

The following officers were elected: Chairman, Professor H. D. Brunk, University of Missouri; Vice-Chairman, Professor J. D. Elder, St. Louis University; Secretary-Treasurer, Miss Mary L. Cummings, University of Missouri.

The following papers were presented:

1. *Some Roman mathematics*, by Professor J. F. Daly, St. Louis University.

From an unpublished manuscript of an unknown author of the late twelfth or early thirteenth century an account is given of how a Roman could find the sum of an arithmetic progression of consecutive numbers and of consecutive even numbers. Also included are the six rules whereby Roman numerals can be multiplied. These rules for multiplication demand the use of an abacus, which is likewise furnished in the text.

2. *The Lebesgue integral for sophomores*, by Professor H. M. MacNeille, Washington University.

The possibility of defining the definite integral in the initial calculus course in a manner that will lead to the Lebesgue rather than the Riemann integral is discussed.

3. *Two-fold generalization of Cauchy's lemma*, by Professor D. E. Coffey, Missouri School of Mines and Metallurgy.

There exist integral solutions x_i of equations

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = a, \qquad x_1 + x_2 + x_3 + x_4 = b,$$

if and only if $a \equiv b \pmod 2$ and $4a - b^2$ is a sum of 3 squares. If $b \geqq 4(1-k)$ and if $b^2 + 2bk + 4k^2 > 3a$, each $x_i > -k$.

4. *A unique construction*, by Mr. H. J. Johnson, Engineer, American Telephone and Telegraph Company, St. Louis.

On one side of a random angle, with a compass lay off points of 1 unit, 2 units and 3 units distance from the vertex of the angle and draw arcs within the angle through these points. Bisect the angle and bisect each half angle limiting the bisecting lines of the half angle to two units radius.

Place a half angle in a new position with its vertex on the intersection of the original bisector with the unit arc keeping the side of the half angle equidistant from the bisecting line. In this position the termini of the sides of the half angle determine two points in free space.

Pivot the compass on the vertex of the original angle, adjust its radius so that a concentric arc may be drawn through these terminal points. This is the unique arc. In case $P=3$, the unique arc is divided at these terminal points into 1/3, 2/3 sections. In case $P=5$, the fifth angle is isolated, *etc.*

5. *Continued fractions, an elementary treatment,* by Mr. C. A. Bridger, Missouri Division of Health, Jefferson City.

Continued fractions have held the attention of mathematicians for centuries. Their properties have aided in the solution of Diophantine equations, in the development of a theory of irrationals, and in the solution of the moment problem in probability, as examples.

Beginning with the long division algorithm, one can develop the c.f. for the rational fraction X in the form

$$a_0+\frac{1}{a_1}+\frac{1}{a_2}+\frac{1}{a_3}\cdots+\frac{1}{a_n},$$

where the a_i are the partial quotients. The convergents to X are P_i/Q_i, where $P_i=P_{i-1}a_{i-1}+P_{i-2}$ and $Q_i=Q_{i-1}a_{i-1}+Q_{i-2}$, $P_0=1$, $Q_0=0$, $P_1=a_0$, $Q_1=1$.

Similar results are derived for the general c.f. constructed from the sequence $\{a_i, b_i\}$ where a_i and b_i are not zero for all i equal to or greater than unity, $a_0=1$, and b_0 may be zero. The transformation for converting a c.f. into an equivalent c.f. is $a_i'=c_{i-1}c_ia_i$ and $b_i'=c_ib_i$, where $c_0=1$ and the other c_i are arbitrary. The convergent 13/8 for the c.f. for the Golden Section is used in art. The c.f. that has only positive integers as its convergents is

$$\frac{1}{1}-\frac{1}{2}-\frac{1}{2}-\frac{1}{2}-\frac{1}{2}\cdots.$$

6. *The teaching of elementary mathematics,* by Professor A. H. Copeland, Sr., University of Michigan. (By invitation.)

C. H. DALTON, *Secretary*

THE MAY MEETING OF THE ALLEGHENY MOUNTAIN SECTION

The thirty-first meeting of the Allegheny Mountain Section of the Mathematical Association of America was held at Westinghouse Research Laboratories, Pittsburgh, Pennsylvania, on May 4, 1957. Dean L. T. Moston, Chairman, presided during the morning and business sessions, and Dr. Morris Ostrofsky presided during the afternoon session. There were 89 persons present, including 52 members of the Association.

During the business session, the following officers were elected for two year terms: Chairman, Professor I. Dee Peters, West Virginia University; Secretary-Treasurer, Dr. B. H. Mount, Westinghouse Electric Corporation, Pittsburgh; Executive Council, Professor J. H. Neelley, Carnegie Institute of Technology, and Professor H. B. Curry, Pennsylvania State University. Also, the Section voted to participate in the national contest program for secondary school students.

The following papers were presented:

1. *On solution of systems of linear equations,* by Dr. A. S. Householder, Army Research Center, Madison, Wisconsin, and Oak Ridge National Laboratories.

Most closed methods of solving linear algebraic equations or inverting matrices can be

classified as methods of factorization and methods of modification. In methods of factorization the matrix is expressed as a product of two matrices, each of which is readily inverted. Generally one of these is triangular, and the other either also triangular or else one with orthogonal rows or orthogonal columns. In methods of modification one proceeds from a matrix of known inverse and modifies an element or group of elements at a time. This method is thought to include Kron's methods of "tearing".

2. *Introduction to electronic digital computers*, by Dr. Ruth O'Donnell Goodman, Westinghouse Research Laboratories, Pittsburgh.

Electronic digital computers, although barely a dozen years old, have already had tremendous influence on the scientific and mathematical world. Each machine is an artfully designed complex of binary decision elements which can carry out arithmetic instructions, making logical decisions during the course of the computation. Since they perform at electronic speeds, even the slowest of these machines can execute hundreds of arithmetic operations per second. When a problem is to be solved by electronic computation, precise instructions for the obtaining of the solution must first be written down, then be translated into the exact "numerical language" which the machine's electronic circuitry can "understand". The preparation of problems for solution by electronic computation is a growing profession for mathematically trained personnel. The remaining papers will outline this work by following through the successive steps which must be taken between the statement of the problem and its final solution.

3. *Mathematical analysis of the problem—the numerical calculation of the Riemann Zeta function near unity*, by Dr. Lowell Schoenfeld, Westinghouse Research Laboratories, introduced by Dr. Ruth O. Goodman.

Consider the calculation, for $s>1$, of

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s} = \sum_{n=1}^{M-1} n^{-s} + \sum_{n=M}^{\infty} n^{-s} \equiv Z_M(s) + R_M(s).$$

In order that the truncation error incurred by dropping $R_M(s)$ should not exceed ϵ, it is necessary that $M \geqq m(s, \epsilon)$. For $s=1.1$ and $\epsilon=10^{-3}$, we have $m(s, \epsilon)=10^{40}$ so that the computation cannot be carried out in this manner by any present or projected computing machine. By using the Euler-Maclaurin sum formula, we find that for each $N \geqq 1$ and some θ_N satisfying $0 \leqq \theta_N \leqq 1$

$$R_M(s) = \frac{1}{M^{S-1}}\left\{\frac{1}{S-1} + \frac{1}{2M} + \sum_{j=1}^{N-1} b_j C_j + \theta_N b_N C_N\right\},$$

where the b_j are modified Bernoulli numbers, $C_1 = s/M^2$, and $C_j/C_{j-1} = (s+2j-3)(s+2j-2)/M^2$ for $j \geqq 2$. It is shown how to determine M and N so as to minimize the work of computation. In particular, if $\epsilon \geqq 10^{-6}$ and $s \leqq 22.1$, then $M \leqq 5$ and $N \leqq 13$, so that, on dropping the term $\theta_N b_N C_N / M^{s-1}$ (whose absolute value now does not exceed ϵ), we have only $(M-1)+1+1+(N-1) = M+N \leqq 18$ terms to compute. The calculation now becomes feasible and much computation time can even be saved for values of s ranging up to 3 and beyond.

4. *Observations on high school mathematics contests*, by Professor E. F. Myers, University of Pittsburgh.

In order to acquaint the members of the Section with the contests held in the (Pittsburgh) area, a report was given on the State Mathematics Tournament in 1955 and a Ninth Grade Algebra Contest, conducted by the Western Pennsylvania Association of the Teachers of Mathematics in five counties in April 1956. Difficulties encountered in administering such tests and tentative plans for future contests by local organizations were presented.

5. *Coordination, entrance conditions and curriculum of high schools*, by Professor J. H. Neelley, Carnegie Institute of Technology.

Mine is a voice "crying in the wilderness", the wilderness of the "wastelands" of American Education. My plea is that we rise up and cut the unimportant and repetitive from our high school mathematics courses and replace them by those topics which lead to a better understanding of college mathematics. Also, that high school mathematics teachers be told what (and in many cases how) to teach these important topics. In consequence, 84 topics desirable for prospective college students were proposed.

6. *Logistics of the computation*, by Dr. R. C. Bollinger, Westinghouse Research Laboratories.

It is the purpose of the talk to try to make clear to those having no previous familiarity with digital computing machines just what one must do to plan a computer program. The various ideas and devices necessary to the planning are illustrated by considering the programming of an actual computer routine to compute values of the Riemann Zeta function. A flow chart which expresses the organization of the computation is constructed in this, the logical analysis, stage of the planning.

7. *Arithmetic, bit by bit*, by Mrs. Aiko Hormann, Westinghouse Research Laboratories, introduced by Dr. R. C. Bollinger.

After a problem is programmed for the machine, it must be further broken down into simple arithmetic operations and then translated into machine code. This translation is comparable to the translation from one human language to another. The similarity and also the difference between the two types of translation are discussed with a few illustrations.

8. *Why doesn't it work?*, by Dr. H. C. Rice, Westinghouse Research Laboratories, introduced by Dr. M. Ostrofsky.

After a suitable emphasis on the inevitability of mistakes in programming a problem for a computer, a survey is made of the kinds of mistakes which can occur, their effect on the behavior of the computer when the program is run, and some of the techniques for finding and correcting them.

I. D. PETERS, *Secretary*

THE MAY MEETING OF THE INDIANA SECTION

The thirty-fourth annual meeting of the Indiana Section of the Mathematical Association of America was held at Purdue University, Lafayette, Indiana on May 11, 1957. Professor C. F. Brumfiel of Ball State Teachers College, Chairman of the Section, presided at both morning and afternoon sessions. There were 77 in attendance including 54 members of the Association.

The following officers were elected: Chairman, Professor C. B. Gass, DePauw University; Vice-Chairman, Professor G. N. Wollan, Purdue University Center, Fort Wayne; Secretary-Treasurer, Professor J. C. Polley, Wabash College.

Chairman Brumfiel announced that the Committee on Awards had awarded three Association medals during the year for high mathematical achievement in the Indiana Science Talent Search.

It was voted that a fall meeting be held this year on October 18 in joint session with the Mathematics Section of the Indiana Academy of Science.

Professor A. W. Tucker of Princeton University, National Lecturer for the Association, gave the invited hour address on the topic, "New Patterns in Mathematical Education."

The following short papers were presented:

1. *Mathematical instruction in Dutch high schools*, by Professor Philip Dwinger, Purdue University.

Some aspects of the teaching of mathematics in Dutch high schools were discussed, such as the subjects taught and the programs for the several final examinations. In addition, attention was called to an important report in 1954 of a committee of the Association of Mathematics Teachers, which, among other things, recommended the introduction of statistics into the program and an intuitive course in plane geometry to precede the regular course in that subject.

2. *The Council, the Association, and the Society*, by Professor C. F. Brumfiel, Ball State Teachers College.

Mathematicians recognize the need for a reorganization of high school mathematics. The content of current texts bears little relation to modern mathematics. These texts abound in gross errors, and an archaic terminology is employed. The organization which sees most clearly the need for the development of a new high school curriculum is the Council, and to effect a change the support of both the Association and the Society is needed. Changes in high school must match the changes that are occurring in colleges and universities. It is to be hoped that these three organizations will, in cooperation, persuade some of the best mathematicians to write texts on the high school level.

3. *Problems of criteria and evaluation*, by Professors M. W. Keller and C. L. Kaller, Purdue University, presented by Professor Keller.

The authors discussed some of the problems inherent in obtaining criteria for admission to and prognosis of success in graduate courses and teaching for mature individuals whose formal training was obtained approximately thirty years ago. A brief report on earlier studies was supplemented by observations based on the current experience of the authors.

4. *Classroom administration*, by Professor G. H. Graves, Valparaiso University.

Since the purposes of the class meeting are to further the student's mastery of material and to increase his ability to gain returns from study, the first essential is to see that his questions are answered, ordinarily by other students. The advantages gained by seat work, board work, and outside work handed in were contrasted. The author felt that note-taking should be discouraged since it distracts from the mental concentration required to take greatest advantage of class work.

5. *Proper cyclic elements, fine cyclic elements, and Lebesgue area*, by Professor C. J. Neugebauer, Purdue University, introduced by the Secretary.

Let Q be a unit square in E_2, and, for (T, Q) a continuous mapping from Q into E_2, let $(T, Q) = lm$, $m: Q \to M$, $l: M \to E_2$ be a monotone-light factorization. For C a proper cyclic element of M, let r_C be the monotone retraction from M onto C. For $L(T, Q)$, the Lebesgue area of (T, Q), the following cyclic additivity formula subsists:

(1) $L(T, Q) = \sum L(r_C m, Q)$, $C \subset M$ (T. Rado, *Length and Area*, Amer. Math. Col. Publ., 30, 1948).

The formula (1) has been generalized and extended by the introduction of a *fine cyclic element* of a mapping (T, J), where J is a closed finitely connected Jordan region (L. Cesari, *Fine cyclic elements of surfaces of the type* γ, Riv. Mat. Univ. Parma). If J is a 2-cell, the fine cyclic elements coincide with proper cyclic elements. In the other cases a fine cyclic element constitutes a suitable decomposition of a proper cyclic element. The above concept of a fine cyclic element can be extended to Peano spaces, and fine cyclic additivity theorems similar to those in paper by E. J. Mickle and T. Rado (*On cyclic additivity theorems*, Trans. Amer. Math. Soc., vol. 66, 1949, pp. 347–365) can be established.

6. *Critical thinking values in introductory modern mathematics*, by Sister Gertrude Marie, Marian College.

Elementary phases of number theory, group theory, the algebra of classes, and modern geometries are cited as source materials for basic experience with definition, undefined terms, relation-

ships expressed in postulates, and theorems resulting from deductive reasoning. The nature of inductive thinking is exemplified by statistical inference. Both induction and deduction are shown to fill important roles in scientific thought, while, in the symbolic formulation of logic, mathematics is identified with critical thinking in its purest interpretation.

7. *Do machines think?*, by Professor R. E. Baer, Purdue University, introduced by Professor Arthur Rosenthal.

Reference is made to papers under similar or related title by Turing, Wilkes, Oettinger, *et al.*, as well as the recent work of Hagelbarger, and Simon and Newell, and that of the Purdue Computation Laboratory. A thinking-like behavior on the part of the universal computer, barring metaphysical but not metamathematical considerations, requires emulation by the machine of both inductive and deductive behavior. The increasing degree of success of machine performance in the two directions is discussed.

8. *Order among complex numbers*, by Mr. Merl Kardatzke, student at Anderson College, introduced by Professor Gloria Olive.

This paper first orders complex numbers by a rule which does not seem to lend itself to a one-to-one correspondence between complex numbers and real numbers. In search for this relationship an analytic expression is found which can order a special set called "semi-countable complex numbers". Finally, binary numbers are used to construct a function which sets up the correspondence which is sought. In conclusion the concept of order is extended to n-dimensional space.

9. *An experiment in teaching calculus over closed-circuit television*, by Professor John Dyer-Bennet, Purdue University, introduced by Professor Arthur Rosenthal.

This paper is a brief report of an experiment conducted at Purdue University, comparing the effectiveness of teaching calculus to small groups over closed-circuit television with that of teaching large groups in lectures. Although the results have not yet been analyzed statistically, they appear to indicate that if effectiveness is measured by the sort of examination commonly used to determine grades, the two methods are about equally good.

10. *The differential*, by Professor H. L. Hunzeker, DePauw University.

The implications arising from the existence of differentials for real functions of one and of several real variables as well as for functions of a complex variable were summarized. An application for the differential of a function of a complex variable was shown in a rather direct proof of the Cauchy Integral Formula.

11. *Some additional remarks on a function defined by means of an infinite radical*, by Professor G. N. Wollan and Mr. D. M. Mesner, Purdue University Center, Fort Wayne, presented by Professor Wollan.

This paper presents some additional properties of the function $f(x)$ defined on $0<x\leqq 1$ by the relation $f(x)=\lim_{n\to\infty} f_n(x)$ where

$$f_1(x)=\sqrt{k+a_1\sqrt{k}},\quad f_2(x)=\sqrt{k+a_1\sqrt{k+a_2\sqrt{k}}},$$

$$f_n(x)=\sqrt{k+a_1\sqrt{k+\cdots+a_n\sqrt{k}}}\ \text{with } n \text{ nested root signs},$$

$n=1, 2, \cdots$, and $a_n=(-1)^{a_n}$ where a_n is the nth digit in the nonterminating binary representation of x. (See this MONTHLY, vol. 63, 1956, p. 614.) The author shows that when $k>2+\sqrt{2}$, although the function has a denumerably infinite set of discontinuities and is not monotone in any subinterval, it is of bounded variation; although it has a value at each point of the interval with $f(x_1)\neq f(x_2)$ when $x_1\neq x_2$, yet the set of values of the function is of measure zero. Furthermore the derivative exists almost everywhere and whenever it exists its value is zero, but there is a non-denumerable set of points (of measure zero) at which the derivative does not exist.

J. C. POLLEY, *Secretary*

THE MAY MEETING OF THE MINNESOTA SECTION

The spring meeting of the Minnesota Section of the Mathematical Association of America was held on May 11, 1957 at Carleton College, Northfield, Minnesota. Professor K. O. May of Carleton College presided at the morning session and Professor Walter Fleming, Chairman of the Section, presided at the afternoon session. There were 72 persons present including 55 members of the Association.

The following officers were elected for 1957–1958: Chairman, Professor O. E. Stanaitis, St. Olaf College; Members of the Executive Committee: Dr. G. E. Baxter, University of Minnesota; Professor Walter Fleming, Hamline University; Mr. F. A. Kros, Remington Rand Univac, St. Paul; Secretary-Treasurer, Professor F. L. Wolf, Carleton College.

It was resolved that the Section express its appreciation to Professor F. C. Smith of the College of St. Thomas for his long and able service as Secretary to the Minnesota Section. This appreciation was expressed by a standing round of applause.

By unanimous vote of the members present the Section resolved to sponsor in the Minnesota region the high school mathematics contest for 1958 as proposed by the National Standing Committee on High School Contests. The chair appointed the following Sectional Committee on High School Contests: G. K. Kalisch (Chairman), F. Hatfield, Donovan Johnson, Walter Fleming, F. L. Wolf.

The following papers were presented:

1. *Trigonometric values that are algebraic numbers*, by Professor K. W. Wegner, Carleton College.

Seventeen equations were presented which have integral coefficients and whose roots include the ninety values $\cos 1°, \cos 2°, \cdots, \cos 89°, \cos 90°$. Ten are of degree less than eight. These ten along with six others that were presented are the only equations with integral coefficients and of degree less than eight with roots that are cosines (or sines) of a rational number of degrees. These sixteen equations were suggested as good ones to use in the algebra and trigonometry classroom.

2. *A derivative relationship*, by Mr. R. P. Winter, College of St. Thomas.

This paper called to attention and reviewed the proof of the interesting derivative relationship $d^2x/dy^2 = -(d^2y/dx^2)(dx/dy)^3$.

3. *Numerical analysis, electronic computers, and information*, by Professor P. C. Hammer, University of Wisconsin. (By invitation.)

In the development of electronic computers some mathematicians have made remarkable contributions while others have held aloof or even actively condemned participants. While the computers cannot live up to all extravagant claims, they do provide mathematicians with means of relieving some of the tedium of their work, and future developments need the guidance of good mathematicians.

The theoretical possibility of duplicating any abstract structure in a computing machine is an indication of a better future for these machines. While application of existing mathematical theories in numerical analysis is by no means complete, more exciting is the fact that numerical analysis has opened new vistas of theoretical mathematics. Specific examples include asymmetric metric functions, spaces of metric functions, and the theory of subadditive functions. Curve fitting is a science in its infancy. The value of computers in general education was discussed.

4. *Remarks on a limit problem*, by Professor O. E. Stanaitis, St. Olaf College.

The limit problem $F(n) = \sum_{p=0}^{N} u_p(n)$, where $u_p(n)$ is a function of n and N is also a function of n that tends to infinity with n was discussed. If $\lim_{n\to\infty} u_p(n) = v_p$ exists, $|u_p(n)| \leqq w_p$, where w_p is independent of n and $p = 0, 1, \cdots$, and $\sum_0^\infty w_p$ converges; then $F(n) \to \sum_{p=0}^{\infty} v_p$ as $n \to \infty$. The theorem is closely related to the M-test for uniform convergence and, in fact, can be considered as a particular case of that test. The superiority of the theorem to the methods of our undergraduate textbooks was demonstrated by examples.

5. *The generalised coconut problem*, by Mr. R. B. Kirchner, Carleton College.

Notions from the calculus of finite differences were applied to the old problem about three sailors with a monkey who gathered a pile of coconuts. The generalization is to n men who divide the pile $n+1$ times, the last division being a division of the remaining coconuts. At each division an arbitrary predetermined number is thrown to the monkey. The sailor then takes one nth of what remains. A solution consists of finding the smallest possible initial amount so that each man receives an integral number of coconuts. A method for solution was given, and it was shown how compact formulas could be derived in special cases.

6. *Completeness and Parseval's equation*, by Professor John M. H. Olmsted, University of Minnesota.

An orthonormal sequence in the space L^2, for an interval $[a, b]$, is called complete in case it is maximal. Equivalent formulations are well known, one being Parseval's equation. If L^2 is replaced by the space RI of Riemann integrable functions on $[a, b]$, these formulations fall into two groups, Parseval's equation implying completeness. Proof that completeness no longer implies Parseval's equation is achieved by extending the space RI to L^2 and showing that the Riemann-integrable functions lying in the orthonormal complement Π of a non-Riemann function are dense in Π.

7. *A minimization problem over a finite set*, by Mr. William J. Hardell, Remington-Rand Univac, St. Paul, Minnesota.

A stepwise procedure for assigning nodes of a finite connected graph G to a finite set of lattice points L of E_n defines a mapping F of G into L. Choice of node and coordinate at each step is such that the maximum length of those edges with one end point placed and the other the node to be placed is a minimum. In the dynamic programming sense F is optimal. Attempts to obtain the mapping of G such that the maximum edge length over all mappings of G is a minimum, by means other than considering all mappings, have been unsuccessful.

F. L. WOLF, *Secretary*

THE MAY MEETING OF THE SOUTHERN CALIFORNIA SECTION

The thirty-seventh regular meeting of the Southern California Section of the Mathematical Association of America was held at San Diego State College, San Diego, California, on May 11, 1957. Professor H. F. Bohnenblust, Chairman of the Section, presided. The attendance was 80, including 63 members of the Association.

At the business meeting it was announced that the mail ballot had resulted in the election of the following officers for the next academic year: Chairman, Professor P. B. Johnson, Occidental College; Vice-Chairman, Professor P. J. Kelly, Santa Barbara College, University of California; Secretary-Treasurer, Mr. R. B. Herrera, Los Angeles City College. These elected officers appointed the following Program Committee: Mr. Clark Lay (Chairman), Pasadena City College; Dr. C. J. A. Halberg, Jr., University of California, Riverside; Professor L. J. Paige, University of California, Los Angeles; and Professor A. L. Whiteman, University of Southern California.

The following program was presented:

1. *Some remarks on Menger's calculus*, by Dr. D. H. Potts, Naval Electronics Laboratory, San Diego.

The three salient features of Menger's presentation of calculus are: (1) his notational innovations, (2) his discussion of *variable* and *function*, and (3) his approach to the problem of the application of calculus to science. These topics are discussed and their basic pedagogic implications brought forth. In particular, the need for notations which clarify concepts yet automatize procedure is stressed. It is becoming increasingly apparent that the student must be taught *mathematics* and not merely mathematical techniques. Yet in introducing rigor we must avoid "rigor mortis". Menger's work is a gigantic step in this direction.

2. *Theory of games and some models of warfare*, by Dr. Melvin Dresher, Rand Corporation, Santa Monica, introduced by the Secretary.

As examples of zero-sum two-person infinite games—*i.e.*, each player has a continuum of strategies—four warfare models were presented:

(1) Defense of targets against attack. Within this model it is shown that only the most valuable targets should be defended. The attacker, however, should use a mixed strategy—concentrate his attack on some one of the defended targets chosen at random subject to a given probability distribution.

(2) Allocation of resources in time. With this model it is shown that during the early stages of a campaign the weaker side should bluff and the stronger side has an optimal pure strategy.

(3) Duels. These analyze the timing of decisions.

(4) Reconnaissance. This evaluates the effect of information on the payoff.

3. *Sets of integers*, by Professor Bodo Volkmann, University of California at Los Angeles, introduced by Professor M. R. Hestenes.

If A and B are infinite sets of nonnegative integers, the sum set $A+B$ is defined as the set of all sums $a+b$, $a\in A$, $b\in B$. A brief survey of some recent results on such sum sets is given, in particular with regard to bases (*i.e.*, sets B for which some "multiple" $hB=B+B+\cdots+B$ contains all nonnegative integers), asymptotic densities of sum sets, and essential components (*i.e.*, sets B satisfying $D(A+B)>D(A)$ for any set A with $0<D(A)<1$, D denoting Schnirelmann density). Some unsolved problems are mentioned, including the following, due to A. Stöhr: Let C_h be the class of bases of order h. For $B\in C_h$, let σ_1, σ_2, and σ_3 denote the $\underline{\lim}$, $\overline{\lim}$, and l.u.b., of $B(n)\cdot n^{-1/h}$, respectively. What are the greatest lower bounds for σ_1, σ_2, and σ_3 as B runs through the class C_h?

4. *Mathematics is changing*, by Professor P. H. Daus, University of California, Los Angeles.

The speaker reviewed the changes that have taken place in the attitude towards mathematics and the contents of various multi-track courses given in the High Schools of California, since the Southern California Section was formed. He also called attention to the changes in the first two years of College Mathematics courses that have already taken place, and the changes that are seriously being considered and are sure to come—for good or ill.

5. *Analytic solutions of interface problems*, by Dr. W. C. Sangren, General Atomic Corporation, San Diego.

By an interface problem is meant a classical boundary value problem with additional internal requirements, or interface conditions, resulting from two or more different internal materials. Many of these problems can be solved analytically by the conventional techniques of the Laplace transform or separation of variables. It is even possible to solve certain problems where the internal boundaries lie along two or more coordinate lines.

6. *Comma-free codes*, by Dr. Basil Gordon, California Institute of Technology.

Let A be an alphabet of n letters, and let S be the set of all k-letter words $(a_1\,a_2\cdots a_k)$, $a_i\in A$. A subset $D\subseteq S$ is called a comma-free dictionary if, whenever $(a_1\cdots a_k)$, $(b_1\cdots b_k)$ are in D, the "overlaps" $(a_2\cdots b_1)$, $\cdots$ $(a_k\cdots b_{k-1})$ are not in D. Let $W_k(n)$ be the greatest number of words such a dictionary can have. The value of $W_k(n)$ is determined for $k=1, 2, 3, 5, 7, 9, 11, 13, 15, 17$, and arbitrary n. It is conjectured that for odd k, $W_k(n)=k^{-1}\sum_{d\mid k}\mu(d)n^{k/d}$, where $\mu(d)$ is the Möbius function. This expression is shown to be an upper bound, but is never attained for k even and $n>3^{k/2}$.

7. *A course in actuarial science*, by Professor Geoffrey Crofts, Occidental College.

An actuary is a scientist and a professional man. As a scientist, he is interested in discovering the pattern of death rates or other rates of decrement. He sets up mathematical models of these rates and draws implications from them. There is now, as always, a need for better models. As a professional man he uses the findings of science and the implications of mathematics to give answers and judgments on matters concerning the operation of pension and insurance funds and related fields.

P. H. DAUS, *Secretary*

THE MAY MEETING OF THE UPPER NEW YORK STATE SECTION

The thirteenth annual meeting of the Upper New York State Section of the Mathematical Association of America was held at Skidmore College, Saratoga Springs, New York, on May 4, 1957. The Chairman of the Section, Professor A. J. Coleman of the University of Toronto, presided at the morning session, and the Vice-Chairman, Professor E. E. Haskins of Clarkson College of Technology presided at the afternoon session. There were 90 persons in attendance, including 56 members of the Association.

At the business meeting the following officers were elected: Chairman, Professor E. E. Haskins, Clarkson College of Technology; Vice-Chairman, Professor Caroline A. Lester, College for Teachers at Albany, State University of New York; Secretary, Professor N. G. Gunderson, University of Rochester.

The Secretary reported that on September 25, 1956, Col. Bessell resigned as Vice-Chairman, and that the Executive Committee then filled the vacancy by choosing Professor E. E. Haskins of Clarkson College of Technology to be Vice-Chairman.

Professor Edith R. Schneckenburger presented the report of the Committee on Mathematics Contests. A resolution was passed to the effect that the Section continue to co-sponsor with the Buffalo Public Schools a contest for the western part of the state, and that the extension of the contest to other areas be investigated.

The appointment of two committees was authorized, one on the strengthening of mathematics in the section, and the other, composed of department chairmen present, to study the specific requirements that have been added by the New York State Education Department for accreditation in the teaching of mathematics in New York State.

The program was as follows:

1. *The training and development of the professional mathematician*, a panel discussion with Professor E. E. Haskins, Clarkson College of Technology, as moderator.

The members of the panel were Mr. W. E. Andrus, Jr., International Business Machines Corporation, Endicott, N. Y., representing industry; Mr. H. E. Webb, Jr., Rome Air Development Center, Rome, N. Y., representing government service; Mr. Frank Hawthorne, New York State Education Department, representing public school teaching; and Professor Paul Olum, Cornell University, representing graduate study.

Mr. Andrus stated that a thorough course in advanced calculus is an absolute "must", and explained how topology is rapidly becoming of greater importance in engineering applications. Mr. Webb spoke of the factors involved in the formulation of the mathematical model representing the problem being dealt with, and especially the importance of probability theory. Mr. Hawthorne reviewed the N. Y. State requirements for high school certification, together with the proposed changes. He was, however, more concerned about the attitude of the teacher toward his subject and teaching, than with the specific preparatory courses. The young teacher should realize that mathematics is a living entity with esthetic qualities of a high order. Professor Olum thought that the undergraduate should have a solid two year course in advanced calculus (analysis), a one year course in modern algebra, and at least an introduction to modern geometry, although he thought that a broad foundation course in modern geometry was probably not very easy to find.

2. *The College Entrance Examination Board's advanced placement program*, by Professor C. R. Keller, College Entrance Examination Board and Department of History, Williams College. (By invitation.)

Under the College Board's Advanced Placement Program able and ambitious students take advanced, college-level work in secondary school followed by advanced placement examinations. These students are then considered by colleges and universities for credit and advanced placement, and many take sophomore courses in college.

Students are challenged in both school and college, school curricula are being thoughtfully revised and standards are being raised, school teachers are being excited by new experiences, college curricula are developing a much needed flexibility, schools and colleges are cooperating in a heartening fashion. For most of these students the Program means educational enrichment rather than acceleration, although acceleration is possible for a few.

In mathematics, the most popular advanced placement subject, advanced placement work means analytic geometry and calculus in the senior year of high school.

3. *The mathematics of "operations research"*, by Dr. A. W. Jones, Bell Telephone Laboratories, New York, New York. (By invitation.)

A brief résumé of the history and nature of operations research was given. Differential and difference equations arising in inventory problems, matrix theory and topology involved in competition models, linear programming, quantity scheduling, probability theory in queuing, renewal, and maintenance problems, symbolic logic in routing and sequence scheduling were given. A new elementary derivation of a linear programming algorithm and a use of complementary solutions in transportation problems were shown to demonstrate the availability of interesting and challenging problems from Operations Research for use in the standard undergraduate and graduate mathematics courses.

4. *A proof of the isoperimetric property of the circle*, by Professor A. G. Davis, Clarkson College of Technology.

The speaker gave a brief review of the history of the isoperimetric problem. The history of integral geometry and the relevant definitions and results were presented. These results were then used to prove the classical isoperimetric inequality that $L^2-4\pi A \geqq 0$. Then it was proved that the equality holds only in the case of a circle. This proof follows the proof given by L. A. Santalo and is of interest because the methods of integral geometry involved use nothing more than elementary calculus.

5. *The matrix equation $X^2=I$ over a finite field*, by Professor J. H. Hodges, University of Buffalo.

Let $GF(q)$, $q=p^f$ where p is prime, denote the finite field of q elements. Consider the problem of finding the number $N(m, x^2-1)$ of $m\times m$ matrices X with elements in $GF(q)$ which satisfy $X^2-I=0$. It can be shown that X is a solution of the equation if and only if X is similar to one of a certain set of canonical matrices. For $p>2$, the canonical matrices are $J_t=\text{diag}\,(I_t, -I_{m-t})$ for $0\leqq t\leqq m$. For $p=2$, they are $H_t=\text{diag}\,(I_{m-2t}, K_1, \cdots, K_t)$ for $0\leqq 2t\leqq m$ and $K_i=\left[\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right]$ for all i. The number $N(m, x^2-1)$ is found by finding the number of distinct X's similar to each canonical matrix and summing over the set of canonical matrices.

6. *Vector methods in elementary analytic geometry*, by Professor Viktors Linis, University of Ottawa.

The author contends that the traditional approach to the analytic geometry with its heavy emphasis on Cartesian coordinate methods and on treatment of conics is of little value from the mathematical and pedagogical point of view. The vector method which introduces such basic concepts as line, parallelism, order, linear dependence, orientation, signed areas and volumes, has the advantages of simplicity, flexibility and applicability. Axiomatic and informal treatments can

be easily blended to suit the needs of students and the temperament of the instructor. The introduced concepts permit natural extension into the theory of linear equations, matrices, transformations, trigonometry, theory of curves and mechanics. The method offers good training in useful formalisms, and at the same time there is no lack of supporting visual interpretations. A course based on the outlined ideas has been prepared by Professor O. Biberstein and has been taught successfully at the University of Ottawa.

7. *Strategy and tactics in mathematics*, by Professor R. B. Davis, Syracuse University.

There are two possible approaches to a mathematical problem; one, by classifying problems and learning specific procedures for their solutions (well-illustrated by many elementary differential equations courses); second, by locating clues within the problem which indicate an appropriate line of attack. This second method, discussed notably by Pólya, involves seeing a problem in such a way that it itself implies a method of attack. What is good or bad about the problem? What is new, or else familiar? One utilizes "approximate "or rough thinking to outline a strategy, then implements this strategy by developing requisite tactical details and devices.

N. G. GUNDERSON, *Secretary*

THE MAY MEETING OF THE WISCONSIN SECTION

The twenty-fifth annual meeting of the Wisconsin Section of the Mathematical Association of America was held at Wisconsin State College, Whitewater, Wisconsin, on May 11, 1957 with Mr. C. J. Vanderlin, Chairman of the Section, presiding. There were 72 persons in attendance, including 41 members.

The following officers were elected for 1957–58; Chairman, Professor R. D. Wagner, University of Wisconsin; Vice-Chairman, Professor John Finch, Beloit College; Secretary-Treasurer, Sister Mary Felice, Mount Mary College.

After a short address of welcome, the following papers were presented:

1. *New patterns in geometry*, by Professor R. H. Bruck, University of Wisconsin.

There is evidence that geometry has almost disappeared from the undergraduate curriculum. Nevertheless, a recent revival in geometric research offers hope of a return to geometry during the next decade. Towards this end, much experimenting along expository lines will be necessary. The present paper represents such an experiment; namely, a partial account of the theory of finite euclidean and projective planes with emphasis on non-standard pictorial representations.

2. *Topology of a line*, by Professor C. E. Burgess, University of Utah, Visiting Lecturer at the University of Wisconsin.

This address consisted of a brief expository discussion of the topological characterization of a line as a connected, separable, linearly ordered space which has neither a first point nor a last point.

3. *Polynomials and Baire's theorem*, by Professor R. P. Boas, Jr., Northwestern University.

The speaker proved the following theorem, discovered by Corominas and Sunyer Balaguer: *If a function $f(x)$ defined on a real interval has derivatives of all orders, and if for each x there is a derivative of some order which vanishes at x, then $f(x)$ is a polynomial.* The proof depends on Baire's category theorem, for which the speaker outlined the necessary preliminaries.

4. *Contest Committee report*, by Professor R. D. Wagner, University of Wisconsin.

Over 1,000 high school students from 132 of Wisconsin's 535 secondary schools took the contest examination which was held on April 6 in twenty cities throughout the state. The top ranking 10 per cent of the students received recognition in the form of certificates, pins, and small cash awards. A set of *World of Mathematics* was awarded to the schools having the 20 top ranking students.

The contest examination was open to any high school students who had completed one year of elementary algebra and was at least taking his second semester of plane geometry. The problems were in part of the short answer type. Over 60 per cent of the examination consisted of essay type questions calling for considerable understanding and ingenuity.

5. *A plea for balance*, by Mr. Arthur Adkins, Wisconsin State Department of Public Instruction, Madison, Wisconsin, introduced by the Secretary.

While recognizing the shortages of trained personnel in mathematics, the sciences, and other technical fields, and admitting the obligation of schools and colleges in helping to meet them, this paper makes a plea for perspective in responding to the pressures on the schools. Feeling that there is a danger, and that there are signs that schools may be yielding to this pressure for more mathematics and science in the curriculum at the expense of other areas and other needs that are also important, the paper asks for balance among specialized fields, and between specialized training and general education, in the public school curriculum.

6. *The new entrance requirements in mathematics of the College of Engineering*, by Professor H. A. Peterson, University of Wisconsin, introduced by the Secretary.

The United States is facing a critical shortage of scientists and engineers. This shortage is not one of *quantity* only, but *quality* as well. It is important that the gifted student be stimulated early and made aware of the opportunities for him in science and engineering. Educators are becoming increasingly aware of the problem and are taking steps to meet the challenge presented. Mindful of the ever increasingly complex problems being encountered by scientists and engineers requiring greater proficiency in the use of mathematics, the College of Engineering of the University of Wisconsin has recently raised the college entrance requirements in mathematics.

7. *Correspondence study courses in relation to the new entrance requirement in mathematics of the College of Engineering*, by Professor H. P. Evans, University of Wisconsin.

The new entrance requirement in mathematics of the University College of Engineering, becoming effective in 1959, includes trigonometry, college algebra, and either solid or analytic geometry. Although most large high schools will adjust to this change in stride, it is anticipated that many small high schools will be unable to offer the necessary courses or to secure adequate staff. Correspondence courses may alleviate these difficulties under a plan which is currently in use and may be expanded. The student carries a correspondence course as part of his high school program and the school system pays the cost of instruction.

8. *Implications for Wisconsin colleges of the 4-yr. mathematics requirement of the University of Wisconsin College of Engineering*, by Dr. L. F. Wahlstrom, Wisconsin State College, Eau Claire.

Seven state and private colleges in Wisconsin brought out the fact that the four-year requirement in high school mathematics which has been adopted by the University of Wisconsin College of Engineering will (1) cause the colleges to re-evaluate their courses and arrange different sequences beginning with calculus and analytic geometry; (2) create a demand for additional upper-level courses in mathematics in the colleges; (3) require better preparation of prospective high school mathematics teachers to teach the mathematics; and (4) probably result in larger enrollments in the upper-level courses.

9. *Implications of the Engineering School's requirement for high schools and report of State Curriculum Committee*, by Mr. W. B. White, North High School, Sheboygan, introduced by the Secretary.

The four-year mathematics requirement for entrance into the University of Wisconsin's School of Engineering, the nationwide emphasis on science and mathematics, and increasing enrollments have brought problems to the state's high schools. Mathematics classes will be of lower aver-

age ability, less experienced teachers may be utilized to teach the additional classes, and there is difficulty in selecting candidates for four years of mathematics at the eighth grade level. A bulletin being prepared by the Statewide Mathematics Curriculum Committee may help solve these problems and will provide suggestions on course content.

SISTER MARY FELICE, *Secretary*

THE JUNE MEETING OF THE PACIFIC NORTHWEST SECTION

The tenth annual meeting of the Pacific Northwest Section of the Mathematical Association of America was held at State College of Washington, Pullman, Washington on June 14, 1957 in conjunction with the 536th meeting of the American Mathematical Society. Professor D. C. Murdoch, Chairman of the Section, presided over the meetings. There were 95 persons in attendance, including 65 members of the Association.

Following a joint dinner meeting with the American Mathematical Society a business meeting was held. The following officers were elected: Chairman, Professor S. G. Hacker, State College of Washington; Vice-Chairman, Professor K. A. Bush, University of Idaho; Secretary-Treasurer, Professor K. S. Ghent, University of Oregon.

After some discussion of the High School Mathematics Contest which the Association plans to sponsor in 1958, the secretary was instructed to report to the national secretary that the contest will be offered in 1958 in British Columbia, Oregon and Washington as in past years, and that there is a possibility that the contest may be offered in Idaho and in Montana. Professor Bush of Idaho and Professor Hurst of Montana State College agreed to investigate the possibility of offering the contest in their states. Professor Murdoch in British Columbia, Professor Chapman at the University of Washington and Professor Ghent in Oregon plan to operate the contests in their respective states as in recent years. Professor Keeping reported that the contest will not be offered in Alberta since the University of Alberta is already sponsoring a provincial contest with the assistance of the Canadian Mathematical Congress.

The afternoon session consisted of an invited address delivered by Professor R. H. Bruck of the University of Wisconsin and a symposium on computing machines. Professor Ostrom introduced Professor Bruck. Professor Lonseth acted as moderator for the symposium.

1. *New patterns in geometry*, by Professor R. H. Bruck, University of Wisconsin. (By invitation.)

See The May Meeting of the Wisconsin Section, abstract 1, p. 627.

2. *Symposium on computing machines*, Moderator, Professor Arvid Lonseth, Oregon State College.

Speakers:

(1) Professor T. E. Hull, University of British Columbia, *The University of British Columbia Computing Center.*

After making plans for more than a year, the University of British Columbia established a computing centre in March, 1957. An Alwac III-E computer is maintained by the Electrical Engineering department. A small staff in the computing centre gives programming courses, develops basic programs and helps users of the equipment. Research projects are being initiated. Besides the need for a more widespread knowledge of computer techniques, there is a need for the modification of existing undergraduate courses and for the development of research in numerical analysis.

(2) Dr. R. E. Gaskell, Boeing Aircraft Company, *Large Scale Industrial Computers and the Universities.*

There is no longer any argument about the fact that the high-speed digital computer has "arrived". The acceptance of the computer as a scientific and business tool is a sign of maturity. Perhaps, then, along with selling the value of the computer, we should be educating potential

consumers of computation in its skillful and economic use. Included in this consumer education might be (1) a notion of the cost of computation, (2) the need for careful preliminary analysis and proper choice of the mathematical model, and (3) the need for complete usage of the computer, for analysis and interpretation as well as for calculation.

(3) Professor O. W. Rechard, State College of Washington, *The computers and the curriculum.*

In this paper the impact of large-scale digital computers on the mathematics curriculum is examined. New courses are called for in programming and coding for digital computers and in the logical theory of automata. Standard courses in numerical analysis and applied mathematics should be adjusted to reflect the changing emphasis in these subjects. Finally, with most of our undergraduate mathematics majors and many of our graduate students taking employment involving the use of digital computers there should be more emphasis on numerical methods in the standard algebra and analysis courses. A computer can make this emphasis feasible and can also serve as an instructional tool to help illuminate certain basic mathematical concepts.

K. S. GHENT, *Secretary*

CALENDAR OF FUTURE MEETINGS

Forty-first Annual Meeting, University of Cincinnati and Hotel Sheraton-Gibson, Cincinnati, Ohio, January 30–31, 1958.

Thirty-ninth Summer Meeting, Massachusetts Institute of Technology, Cambridge, Massachusetts, August 25–28, 1958.

The following is a list of the Sections of the Association with dates of future meetings, so far as they have been reported to the Associate Secretary.

ALLEGHENY MOUNTAIN, Washington and Jefferson College, Washington, Pennsylvania, May, 1958.

ILLINOIS, Illinois College, Jacksonville, May 9–10, 1958.

INDIANA, DePauw University, Greencastle, October 18, 1957.

IOWA, Drake University, Des Moines, April 18, 1958.

KANSAS

KENTUCKY, University of Kentucky, Lexington, April, 1958.

LOUISIANA-MISSISSIPPI, Loyola University, New Orleans, February 21–22, 1958.

MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA, Georgetown University, Washington, D. C., December 7, 1957.

METROPOLITAN NEW YORK

MICHIGAN, University of Michigan, Ann Arbor, March, 1958.

MINNESOTA, State Teachers College, Mankato, October 5, 1957.

MISSOURI, University of Missouri, Columbia, Spring, 1958.

NEBRASKA, University of Nebraska, Lincoln, April 19, 1958.

NEW JERSEY, Fairleigh Dickinson University, Rutherford, November 2, 1957.

NORTHEASTERN, Dartmouth College, Hanover, New Hampshire, November 30, 1957.

NORTHERN CALIFORNIA, San Francisco State College, January 18, 1958.

OHIO, Denison University, Granville, April, 1958.

OKLAHOMA, Oklahoma City University, October 25, 1957.

PACIFIC NORTHWEST, Oregon State College, Corvallis, June 20, 1958.

PHILADELPHIA, Haverford College, Haverford, Pennsylvania, November 30, 1957.

ROCKY MOUNTAIN, Colorado State College, Greeley, Spring, 1958.

SOUTHEASTERN, University of Florida, Gainesville, March 14–15, 1958.

SOUTHERN CALIFORNIA, Pasadena City College, March 8, 1958.

SOUTHWESTERN, University of New Mexico, Albuquerque, April 11–12, 1958.

TEXAS, Baylor University, Waco, April, 1958.

UPPER NEW YORK STATE, École Polytechnique and University of Montreal, Montreal, Quebec, Canada, May, 1958.

WISCONSIN, Carroll College, Waukesha, May, 1958.

For College Courses . . .

SOLID GEOMETRY

HUGO MANDELBAUM *and* **SAMUEL CONTE**
—both Wayne State University

This new textbook offers a sound introduction to modern mathematical thinking consistently applied to the teaching of elementary solid geometry. Concepts of projective geometry are presented to break the limitations imposed on Euclidean geometry. These concepts allow the student to view cylinders and prisms, cones and pyramids as interrelated by their mode of generation. *296 ills., tables; 261 pp.*

Outstanding Teachability:

Emphasis on understanding and application rather than axiomatics and systematic proofs of theorems

Numerous good-sized, clear illustrative figures—many of them shaded to facilitate space conception—accompany the text, examples, and exercises

TRIGONOMETRY

ROY DUBISCH, *Fresno State College*

Relating trigonometry more closely to other mathematics courses, this popular book presents trigonometric functions as functions of real numbers. Trigonometric functions of angles are a supporting topic. Book emphasizes the distinction between a function and a function value. Provides many examples and problems. Arc length protractor and scale included.
119 ills., tables; 396 pp.

Points of Interest:

Stresses those procedures and problems which are most generally useful in applications of trigonometry

Addition formulas are derived in full generality by a simple application of the distance formula of analytic geometry set forth in the first chapter

THE RONALD PRESS COMPANY • 15 East 26th St., New York 10

DYNAMIC PROGRAMMING

By **RICHARD BELLMAN**

The theory of dynamic programming, rapidly becoming more widely known and used in problems involving multi-stage decision processes, is presented from the beginning by the scientist who coined the term and developed the theory in its early stages. In introducing this theory, Richard Bellman furnishes a new and versatile mathematical tool for the treatment of many novel and interesting problems. Since the methods and application are both new, the book is written at a moderate mathematical level which does not obscure the simplicity of the guiding ideas. Many examples are provided to help the reader in mastering the material. ***A RAND Corporation Research Study.***

330 pages. $6.75

Order from your bookstore, or

PRINCETON UNIVERSITY PRESS • Princeton, N.J.

OPERATIONS RESEARCH ANALYST

We are seeking a particularly creative individual with a Ph.D. or M.S. and equivalent experience, possessing research experience in Mathematics or Mathematical Statistics plus practical knowledge of high speed computers to develop analytical methods and models for solution of military and non-military Operations Research problems for a challenging assignment in our Operations Research Group.

Current activities of this Group include:

1. A study of digital computer techniques for simulating air traffic and air traffic control procedures.
2. An investigation of properties of communication networks through digital computer simulation of systems of interest.
3. Research concerned with methods for prediction of traffic distribution over road networks.
4. Development and evaluation of optimum path selection techniques for finding the shortest or least cost path through a maze.
5. Research studies of data processing requirements and techniques for major commercial and military operations.
6. Analysis of hospital service operator functions to improve doctor paging and sign-in procedures.
7. Multivariate statistical analysis of physical and biological systems, with particular emphasis on design of experiments.

This position offers unusual freedom of activity and diversified opportunity to the man who wishes to fully utilize his creative ability.

As a leading independent research organization, we offer competitive salaries, excellent benefits and generous relocation allowance.

Send complete resume to:

A. J. Panoral
Armour Research Foundation
of Illinois Institute of Technology
10 West 35th Street
Chicago 16, Illinois

Towards a more integrated Mathematics Program

Selections From
MODERN ABSTRACT ALGEBRA

Richard V. Andree

This soundly-conceived new text is a detailed study of Abstract Algebra which covers a wide range of topics, including Boolean Algebra, Group Theory, Linear systems, and other Algebraic varieties and ideals. Particularly suitable for the new two-hour mathematics course now being offered in a number of schools.

1958

FUNDAMENTAL MATHEMATICS

Leslie H. Miller

Offering a sound and thoroughly tested course for any student inadequately prepared for college mathematics, this combination text and workbook is effectively geared for use in Freshman-level courses, Remedial Mathematics, or as a supplementary text for dealing with individual weaknesses.

1957

other titles

ANALYTICAL GEOMETRY, Third Edition
1955
Charles H. Sisam
William F. Atchinson

FUNDAMENTALS OF COLLEGE MATHEMATICS
1954
John C. Brixey
Richard V. Andree

MODERN TRIGONOMETRY
1955
John C. Brixey
Richard V. Andree

FIRST COURSE IN ALGEBRA FOR COLLEGES
1955
Lovincy J. Adams

Henry Holt and Co., 383 Madison Ave., New York 17, N.Y.

APPLIED MATHEMATICIANS

Exceptional opportunities are now available for applied mathematicians at the Jet Propulsion Laboratory. Qualified people are needed for advanced research in numerical analysis and the application of automatic digital computers to the solution of problems in applied mathematics.

A Ph.D. degree, or its equivalent, with a background in numerical analysis is required. Ability to work independently is desirable. The work is supported by modern digital computer facility and excellent staff accommodations.

The Laboratory is a continuing operation devoted to scientific research and development and offers many challenging opportunities for increasing responsibility in an expanding activity.

Your resume will receive immediate consideration if forwarded as soon as possible.

jpl

JET PROPULSION LABORATORY

A DIVISION OF CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

MATHEMATICIANS for IBM

IBM Research Laboratories at Poughkeepsie and Yorktown, N.Y., offer exceptional career opportunities to a select group of highly qualified mathematicians with an M.S. or Ph.D. degree. Research includes investigation of new methods for programming data, planning of machine systems, new principles for electronic data processing machines and giant computers. This represents an unusual opportunity to be associated with IBM in an important research capacity. Experience in one, or several, of the following types of work is preferred:

Statistical programming
Numerical analysis
Network theory
Character recognition

IBM'S ADVANTAGES

- Opportunities for advanced training, publication of papers and independent research.
- "Small team" work system and advancement based on merit.
- Excellent record of stability and growth.
- Numerous company-paid benefits.
- Relocation expenses paid to Poughkeepsie and Yorktown, N.Y.

HOW TO APPLY

Please write, outlining your qualifications and experience, to:

Mr. G. W. Smith
IBM Corporation, Dept. 1410
Box 218
Yorktown Heights, N.Y.

INTERNATIONAL BUSINESS MACHINES CORPORATION

Medical Expenses mount up fast

Now, through TIAA's new MAJOR MEDICAL EXPENSE INSURANCE, colleges can help free their staff members from concern over the financial problems of medical care for themselves and their families.

TIAA MAJOR MEDICAL is issued on the Group basis and can be added to a "base plan" such as Blue Cross-Blue Shield or designed to provide the entire medical expense insurance program at the college.

Colleges, universities, independent schools, and certain foundations and other nonprofit educational or research institutions are eligible whether or not they now have a TIAA retirement or insurance program.

To learn more about this important protection, complete and return:

Teachers Insurance and Annuity Association
522 Fifth Avenue, New York 36, New York

Please send information on MAJOR MEDICAL EXPENSE INSURANCE to:

Name_____

Title_____

Employing Institution_____

Address_____

F

Fresh & Exciting

◆

for mathematics majors and those preparing to teach

◆

in press

The Foundations & Fundamental Concepts of Mathematics

by HOWARD EVES & CARROLL V. NEWSOM

The teaching experience and exceptional writing skill of two noted mathematicians are evident throughout this marvelously clear, interesting modern treatment of the meaning and nature of mathematics—the first modern treatment suitable for undergraduate as well as graduate students.

just published

Understanding Arithmetic

by ROBERT L. SWAIN

Already proving a stimulating and enlightening new adventure into mathematics for hundreds of students, particularly those preparing to teach. Outstanding in its treatment of the theory of numbers, meaning of units, the accuracy concept, and other basic concepts.

Class-tested

◆

for basic courses

◆

Calculus

by JACK R. BRITTON

The preferred text of teachers across the country for its exceptionally clear, logical, easy-to-teach organization and wealth of excellent problems.

Algebra for College Students

REVISED

by JACK R. BRITTON & L. CLIFTON SNIVELY

Used in over 80 leading colleges and universities, this text has become a stand-by for the basic college algebra course, especially liked for its effectiveness with the poorly prepared student.

R I N E H A R T

& COMPANY, INC. 232 Madison Ave., New York 16

Important McGraw-Hill Books

APPLIED MATHEMATICS FOR ENGINEERS & PHYSICISTS

By Louis A. Pipes, University of California at Los Angeles. New Second Edition. Ready for Second Semester Classes.

Written especially for the general advanced math course, this new second edition aims to provide the engineer and applied physicist with the principal mathematical techniques necessary for the analysis of the usual problems that arise in practice. It covers those topics which are essential for an understanding of present-day engineering analysis i.e.—infinite series, complex numbers and the complex variable, Fourier equations, modern algebraic methods and ordinary nonlinear differential equations.

LINEAR PROGRAMMING AND ECONOMIC ANALYSIS

By Robert Dorfman, Harvard University; Paul A. Samuelson and Robert M. Solow, both of the Massachusetts Institute of Technology. Ready in December.

With emphasis on economic aspects of linear programming, this new mathematical economics text authoritatively and definitively shows its relationship to standard economic analysis. Higher mathematics, avoided as far as possible, is explained wherever used. Recent advances in linear programming, activity analysis, dynamic programming, welfare economics, simplex and other computational methods are included.

MATHEMATICS FOR SCIENCE AND ENGINEERING

By Phillip A. Alger, The General Electric Company. 384 pages, $6.95 (text edition available)

The author's purpose is to help the student or engineer to use mathematics effectively, by showing the unity and simplicity of the basic mathematical ideas and by making numerical calculation easier. Beginning with a quick review of arithmetic, the book covers algebra, trigonometry, calculus, probability, and lineal differential equations, giving special emphasis to complex numbers, infinite series, and methods of approximation. Here are almost all the mathematical procedures and methods an engineer will need in his career.

MODERN GEOMETRY: An Integrated First Course

By Claire Fisher Adler, New York University. Ready in December

This text for both math majors and prospective secondary school teachers unifies a vast amount of accumulated material in foundations of geometry and in the fields of Euclidean, projective, and Non-Euclidean geometry. Here is a clear and understandable book that bridges the gap between Euclidean and the modern abstract theories. Historical backgrounds accompany each new development to acquaint the reader with the momentous discoveries of the past 200 years.

Send for copies on approval

McGRAW-HILL BOOK COMPANY

330 West 42nd Street • New York 36, N. Y.

ANALYTIC GEOMETRY, Fifth Edition

by CLYDE E. LOVE *and* EARL D. RAINVILLE

Flexibility of presentation characterizes this text, making it adaptable to the requirements of individual instructors. This fifth edition includes: circles of Appolonius, the distance formulas in polar coordinates, radical axis, common chord, tangents to a conic from an external point, and chord of contact. There are over 2000 carefully-constructed exercises. *1955, 302 pages, $4.25*

DIFFERENTIAL AND INTEGRAL CALCULUS Fifth Edition

by CLYDE E. LOVE *and* EARL D. RAINVILLE

Retaining the high standard of the previous editions, this fifth edition incorporates additional material. Among the added topics are: circle of curvature, integral test for infinite series, summation of power series, oblique and curvilinear asymptotes, evaluation of iterated integrals by change of coordinate system, and a short appendix on rigorous presentation of limits. *1954, 526 pages, $4.95*

VECTOR AND TENSOR ANALYSIS

by NATHANIEL COBURN

This book presents a significant treatment of vector and tensor analysis and shows the value of these subjects as tools in studying other topics such as: rigid-body dynamics, perfect fluids, differential geometry, finite deformation in elasticity, viscous fluids, compressible fluids and turbulence. Vector analysis is treated in the conventional Gibbs manner, i.e. by use of dissected line segments.

1955, 341 pages, $7.00

THEORY OF NUMBERS

by B. M. STEWART

With 33 lessons covering the basic topics for the first course in number theory and an excellent selection of supplementary topics, this text is intended for the one-semester course. Among the advanced topics covered are: equivalence relations, abstract mathematical systems, groups of transformation, matrices, domains, and fields. A number of unusual and stimulating examples, figures and quotations are incorporated into the text. *1952, 261 pages, $5.75*

The Macmillan Company

60 FIFTH AVENUE, NEW YORK 11

GEORGE BANTA COMPANY, INC., MENASHA, WISCONSIN, U.S.A.

TO LESTER R. FORD ON HIS SEVENTIETH BIRTHDAY

The Sixth
HERBERT ELLSWORTH SLAUGHT
MEMORIAL PAPER

Published as a supplement to the AMERICAN MATHEMATICAL MONTHLY
Volume 64 Number 8
October, 1957

DEDICATION

October 25, 1956, was the seventieth birthday of Lester R. Ford, President of the Mathematical Association of America (1947–1948), editor of the AMERICAN MATHEMATICAL MONTHLY (1942–1948). In the spring of 1956, some of the numerous friends and former students of Dr. Ford decided to dedicate papers to him on this occasion as tokens of their appreciation and friendship. The collection of manuscripts was presented to Dr. Ford on his birthday by Drs. W. L. Duren, Jr., Karl Menger, and G. T. Whyburn. Some of the papers are related to the fields of Ford's major interests: complex functions, interpolation, differential equations, and numerical analysis. Other papers were inspired by remarks that Ford made in talks and lectures. Still others have to do with Ford's former activity as editor of the MONTHLY, where he started the series of papers with the titles "What is . . . ?"

In this number of the Slaught papers, the papers dedicated to Dr. Ford are arranged alphabetically according to the authors' names. All authors are united in the feeling expressed in the final remark contained in the paper of G. T. Whyburn.

KARL MENGER
GORDON PALL

TABLE OF CONTENTS

Dedication iii

Distance Geometry of Metric Arcs . . . . . WILLIAM R. ABEL AND LEONARD M. BLUMENTHAL 1

The Approximation of Logarithmic Numbers . . . H. T. DAVIS 11

Mathematical Induction in Sets . . . . . W. L. DUREN, JR. 19

Calculation of Moments for a Cantor-Vitali Function . G. C. EVANS 22

Solution of a Ranking Problem from Binary Comparisons . . . . . L. R. FORD, JR. 28

On the Computation of Certain Roots by the Use of the Binomial Series . . . . . JOHN W. GREEN 34

An Excerpt from the Works of Euler . . . RUDOLPH E. LANGER 37

Quantum Mechanics and Hilbert Space . . . GEORGE W. MACKEY 45

Multiderivatives and Multi-Integrals . . . . KARL MENGER 58

Generalized Jacobi Expansions and Corresponding Derivatives . . . . . GORDON PALL 71

On the Space of Oriented Lines in Euclidean Three-Space . . . . . TIBOR RADÓ 79

A Covering Theorem for Univalent Functions . . . W. T. SCOTT 90

Some Convergence Problems for Continued Fractions . H. S. WALL 95

Sense and Orientation on the Disk . . . . . G. T. WHYBURN 103

DISTANCE GEOMETRY OF METRIC ARCS*

WILLIAM R. ABEL and LEONARD M. BLUMENTHAL,
University of Missouri

1. Introduction. This paper, the first of a series, investigates metric arcs that possess at each point a finite Menger curvature. The properties of this class of arcs that are established here center about those finite subsets of an arc that form *lattices* and *homogeneous* chains. It is shown, for example, that for all but a finite number of positive integers n, each n-lattice L_n of an arc A (with finite Menger curvature at each point) is (i) unique, and (ii) a homogeneous $\lambda(n)$-chain, where $\lambda(n)$ is the distance of two consecutive points of L_n. The concept of *metrically monotone arcs* is introduced in Section 5, and it is seen that each metric arc with everywhere finite Menger curvature is the sum of a finite number of such arcs.

Other theorems established are (1) the length of *any* metric arc (rectifiable or not) is the *limit* of the lengths of inscribed n-lattices (a result that is not valid for semimetric arcs) and (2) every metric ptolemaic arc that is a geodesic (that is, *locally*, a metric segment) is a metric segment. These two results fill lacunae in the proof of the fundamental theorem that characterizes metric segments as arcs with everywhere vanishing Menger curvature which Schoenberg obtained by extending Menger's procedure for euclidean arcs to a more general environment. It is indicated, finally, how our study leads to a new approach to that characterization theorem.†

2. Definitions and preliminary results. A metric arc is a homeomorph of a line segment, with a distance defined that satisfies the postulates for a metric space. Denoting such an arc, with initial point a and terminal point b by $A(a, b)$, a finite subset $P=(p_1, p_2, \cdots, p_m)$ is *normally ordered* provided these points are encountered in the order of their subscripts as $A(a, b)$, is traversed from a to b. The length $L[A]$ is the least upper bound of the numbers $L[P]=p_1p_2+p_2p_3+\cdots+p_{m-1}p_m$ for all normally ordered subsets $P=(p_1, p_2, \cdots, p_m)$ of $A(a, b)$.

A normally ordered subset P of $A(a, b)$ is a homogeneous ϵ-chain, $\epsilon>0$, provided (1) $ap_1<\epsilon$, $p_mb<\epsilon$, (2) $p_ip_j=\epsilon$ for $|i-j|=1$, $(i, j=1, 2, \cdots, m)$ and (3) $p_ip_j\geqq\epsilon$, for $|i-j|>1$, $(i, j=1, 2, \cdots, m)$.

It is easily proved that for each positive ϵ, each metric arc contains a homogeneous ϵ-chain.‡

A normally ordered $(n+1)$-tuple $p_0, p_1, \cdots, p_n$ of $A(a, b)$ forms an n-lattice provided $p_0=a$, $p_n=b$, and $p_0p_1=p_1p_2=\cdots=p_{n-1}p_n$.

An elegant proof by Schoenberg shows that for every positive integer n, every continuous curve with distinct endpoints, in any metric space (indeed in

* Presented to the American Mathematical Society, October 27, 1956.

† For references to the literature, as well as to those concepts of distance geometry not defined in this paper, see L. M. Blumenthal, Theory and Applications of Distance Geometry, The Clarendon Press, Oxford, 1953, referred to in the following footnotes as Distance Geometry.

‡ Distance Geometry, pp. 60–61.

any semimetric space with continuous distance function) contains an n-lattice.† We shall refer to this important result as the *n-lattice theorem*. The distance p_ip_{i+1} of consecutive points p_i, p_{i+1} $(i=0, 1, \cdots, n-1)$ of an n-lattice is denoted by $\lambda(n)$.

The distinction between an n-lattice and a homogeneous chain is clear. Apart from the fact that the initial and terminal points of a homogeneous chain need not coincide with the initial and terminal points of the arc (while this is required for an n-lattice) it is important to observe that condition (3) for a homogeneous chain has no counterpart in an n-lattice. It should also be remarked that neither homogeneous ϵ-chains nor n-lattices are necessarily uniquely determined by the given ϵ and n, respectively.

If $p\in A(a, b)$, the metrization of the curvature $K(p)$ of the arc at p, due to Menger, is expressed by

$$K(p) = \lim_{q,r,s\to p} \frac{\{-D(q, r, s)\}^{1/2}}{qr\cdot rs\cdot qs}, \qquad q\neq r\neq s\neq q,$$

where $q, r, s,\in A(a, b)$ and

$$D(q, r, s) = (qr + rs + qs)(qr + rs - qs)(qr - rs + qs)(qr - rs - qs).‡$$

The expression whose limit is taken is called the *curvature of the point triple* q, r, s and is denoted by $K(q, r, s)$. It is seen that $K(q, r, s)$ is the reciprocal of the radius (perhaps infinite) of a circle (in the euclidean plane) that contains three points q', r', s' with $qr=q'r'$, $rs=r's'$, $qs=q's'$. Clearly, $K(q, r, s)=0$ if and only if q', r', s' are on a line.

3. Lattices and arc length. The principal result of this section exhibits arc length as the limit of the lengths of n-lattices inscribed in the arc. This extension to metric arcs of a classical property of euclidean arcs is very useful in our study. It is observed that the result is not valid in general semimetric spaces with continuous distance function.§

LEMMA 3.1. *If A is any metric arc and $\lambda(n)$ is the side of any n-lattice of A, then* $\lim_{n\to\infty}\lambda(n)=0$.

Proof. Let f denote a homeomorphism from the unit interval $I=[0, 1]$ to the arc A. Since f is uniformly continuous in I, there corresponds to each positive ϵ a positive δ such that if $p, q\in I$ and $pq<\delta$, then $f(p)f(q)<\epsilon$. Now if N is a positive integer such that $N\delta>1$, then for every $n>N$, each n points of I has at

† Distance Geometry, pp. 73–74.

‡ Distance Geometry, p. 75.

§ Consider, for example, the semimetric space (an arc) with continuous distance function obtained by squaring the euclidean metric of the segment $[0, 1]$. For every positive integer n, the points $0, 1/n, 2/n, \cdots, (n-1)/n, 1$ form an n-lattice of the arc, but $\lim_{n\to\infty} n\cdot\lambda(n)=\lim_{n\to\infty} 1/n=0$. Note that $\lim_{n\to\infty} n^2\cdot\lambda(n)=1=$l.u.b. $L[P]$, for all finite normally ordered subsets P of the arc; that is, $\lim_{n\to\infty} n^3\cdot\lambda(n)=L(\text{arc})$. This suggests associating an integer k with a semimetric space provided $L[A]=\lim_{n\to\infty} n^k\cdot\lambda(n)$ for each arc A of the space.

least one consecutive pair with distance less than δ. Consequently, the side of any n-lattice of A, with $n>N$, is less than ϵ and so $\lim_{n\to\infty}\lambda(n)=0$.

LEMMA 3.2. *Let C denote a constant and Q any normally ordered subset of a metric arc $A(a, b)$ with $L[Q]>C$. There exists a positive integer N such that for every $n>N$, $n\cdot\lambda(n)>C$, where $\lambda(n)$ is the side of any n-lattice of $A(a, b)$.*

Proof. Let $Q=\{q_0, q_1, \cdots, q_m\}$ and consider the point q_i, where i is an arbitrary but fixed one of the indices $0, 1, \cdots, m$. We show that corresponding to each $\sigma>0$ there is a positive integer N_i^* such that for every $n>N_i^*$ each n-lattice of $A(a, b)$ has a point on $A(q_i, b)$ (the sub-arc of $A(a, b)$ with endpoints q_i, b) with distance from q_i less than σ. If $i=0$ and $q_0=a$ then a is itself such a point, while b serves in case $i=m$ and $q_m=b$. Supposing that q_i is neither a nor b, consider the two subsets $A(a, q_i)$ and $A(q_i, b)-U(q_i; \sigma)\cdot A(q_i, b)$ of $A(a, b)$, where $U(q_i; \sigma)$ is the (open) spherical neighborhood of q_i with radius σ. If $A(q_i, b)-U(q_i; \sigma)\cdot A(q_i, b)=0$, then b again serves as the desired point; in the contrary case, the two closed and compact sets have a positive distance δ.

Since $\lim_{n\to\infty}\lambda(n)=0$, there exists a positive integer N_i^* such that $\lambda(n)<\delta$ whenever $n>N_i^*$. Clearly, for each n-lattice L_n, each of the sets $L_n\cdot A(a, q_i)$, $L_n\cdot[A(q_i, b)-U(q_i; \sigma)A(q_i, b)]$ is non-null, and consequently for every $n>N_i^*$, every n-lattice of $A(a, b)$ contains a point with the desired property. In a similar manner, to each $\sigma>0$ there is a positive integer N_i^{**} such that every n-lattice with $n>N_i^{**}$ has a point on $A(a, q_i)$ with distance from q_i less than σ.

Now if $n>N(\sigma)=\max[N_0^*, N_0^{**}, N_1^*, N_1^{**}, \cdots, N_m^*, N_m^{**}]$, then every n-lattice L_n has points on each side of q_i with distances from q_i less than σ. Let the first such point on $A(q_i, b)$ be labelled $q_{j_i}^*$, and the last such point on $A(a, q_{i+1})$ be denoted by $q_{j_{i+1}}^{**}$ ($q_{j_i}^*$ is the first point of L_n to the right of q_i, with distance from q_i less than σ, and $q_{j_{i+1}}^{**}$ is the last point of L_n to the left of q_{i+1}, with distance from q_{i+1} less than σ). Then $q_iq_{i+1}\leqq q_iq_{j_i}^*+(\text{length of } L_n \text{ from } q_{j_i}^* \text{ to } q_{j_{i+1}}^{**})+q_{j_{i+1}}^{**}q_{i+1}$; that is,

$$q_iq_{i+1} < 2\sigma + (\text{length of } L_n \text{ from } q_{j_i}^* \text{ to } q_{j_{i+1}}^{**}).$$

It follows that for every $n>N(\sigma)$, and each n-lattice L_n

$$n\cdot\lambda(n) + 2m\sigma > L[Q] > C.$$

Take $\sigma=(L[Q]-C)/2m$ and put $N=N(\sigma)$. Then for $n>N$, $n\cdot\lambda(n)+(L[Q]-C)>L[Q]$; that is, $n\cdot\lambda(n)>C$, and the lemma is proved.

THEOREM 3.1. *If A is any metric arc, then $L[A]=\lim_{n\to\infty} n\cdot\lambda(n)$.*

Proof. If $L[A]$ is finite then to each $\epsilon>0$ there is a normally ordered subset Q of A such that $L[Q]>L[A]-\epsilon$. Letting $L[A]-\epsilon$ be the constant C of the above lemma, there exists a positive integer N such that for every n-lattice with $n>N$, $L[A]\geqq n\cdot\lambda(n)>L[A]-\epsilon$; that is, $L[A]=\lim_{n\to\infty} n\cdot\lambda(n)$.

If $L[A]$ is infinite, for every positive constant C there exists a normally

ordered subset Q of A with $L[Q]>C$. Then by Lemma 3.2, there exists a positive integer N such that for $n>N$, the length of every n-lattice of A exceeds C. Hence $\lim_{n\to\infty} n\cdot\lambda(n)=\infty=L[A]$.

Remark. It may be of interest to note that if A is a metric arc and Q is a finite normally ordered subset of A, there might not exist any n-lattice of A such that $n\cdot\lambda(n)\geqq L[Q]$. For let a, b, c be three non-collinear points of the euclidean plane with ac/bc irrational. Then no n-lattice of the arc $A(a, b)=\text{seg}\,[a, c]+\text{seg}\,[c, b]$ contains the point c and consequently for every n-lattice of $A(a, b)$, $n\cdot\lambda(n)<L[Q]=ac+cb$, where $Q=\{q_0, q_1, q_2\}$, $q_0=a$, $q_1=c$, $q_2=b$.

4. Lattices and homogeneous chains. The principal results of this section, as well as Section 5, are established with the aid of the following lemma.

LEMMA 4.1. *If a metric arc A has finite Menger curvature at each point, then a positive number ϵ exists such that for every point p of A, r, s, $t\in U(p;\epsilon)\cdot A$, $rs=st$ imply $rs<rt$.*

Proof. If the lemma is false then for each positive integer n there exists a point p_n of A and three points r_n, s_n, t_n of $U(p;\epsilon_n)\cdot A$ with $r_ns_n=s_nt_n\geqq r_nt_n$, where $\{\epsilon_n\}$, $(n=1, 2, \cdots)$, is an infinite sequence with limit zero. Since A is closed and compact, there is a point p of A such that $\lim_{i\to\infty} p_{n_i}=p$, where $\{p_{n_i}\}$ is a subsequence of $\{p_n\}$. Clearly $\lim_{i\to\infty} r_{n_i}=\lim_{i\to\infty} s_{n_i}=\lim_{i\to\infty} t_{n_i}=p$.

Now

$$K^2(r_{n_i}, s_{n_i}, t_{n_i}) = [4(r_{n_i}s_{n_i})^2 - (r_{n_i}t_{n_i})^2]/(r_{n_i}s_{n_i})^4 \geqq 3/(r_{n_i}s_{n_i})^2,$$

and consequently

$$K^2(p) = \lim_{i\to\infty} K^2(r_{n_i}, s_{n_i}, t_{n_i}) = \infty,$$

in contradiction to the assumed finiteness of the curvature of A at each point.

THEOREM 4.1. *If a metric arc $A(a, b)$ has finite Menger curvature at each point, then a positive integer N exists such that for $n>N$ every n-lattice of $A(a, b)$ is a homogeneous $\lambda(n)$-chain (that is, "almost all" n-lattices of the arc are homogeneous chains).*

Proof. According to Lemma 3.1, a positive integer N exists such that for $n>N$, $\lambda(n)<\min(\epsilon, ab)$, where ϵ is the epsilon of the preceding lemma, and $\lambda(n)$ is the side of any n-lattice $L_n=(p_0^n, p_1^n, \cdots, p_n^n)$ of $A(a, b)$. To prove that L_n is a homogeneous $\lambda(n)$-chain, it suffices to show that $p_i^np_{i+j}^n\geqq\lambda(n)$, $(i, j=0, 1, \cdots, n;\ 0<i+j\leqq n)$.

Suppose two distinct points p_i^n, p_{i+j}^n of L_n exist such that $p_i^np_{i+j}^n<\lambda(n)$. Then $j>1$ and the sub-arc $A(p_{i+1}^n, b)$ contains p_{i+j}^n. Let $f(p_i^n)$ denote a foot of p_i^n on this sub-arc.

Case 1. *The point $f(p_i^n)$ is distinct from b.* Since $p_i^np_{i+1}^n=\lambda(n)>p_i^np_{i+j}^n$, then $f(p_i^n)\neq p_{i+1}^n$ and $p_i^nf(p_i^n)\leqq p_i^np_{i+1}^n=\lambda(n)<\min(\epsilon, ab)$. For each positive integer k

select points x_k, y_k such that $x_k \in A(p_{i+1}^n, f(p_i^n))$, $y_k \in A(f(p_i^n), b)$, $0 < x_k f(p_i^n) = y_k f(p_i^n) < 1/k$. If $x_k p_i^n \leqq y_k p_i^n$ put $x_k' = x_k$ and choose $y_k' \in A(f(p_i^n), y_k)$ so that $y_k' p_i^n = x_k' p_i^n$; similarly, if $x_k p_i^n > y_k p_i^n$, define $y_k' = y_k$ and select a point x_k' of $A(x_k, f(p_i^n))$ such that $x_k' p_i^n = y_k' p_i^n$. It is easily seen that such selections are possible, and so two infinite sequences $\{x_k'\}$, $\{y_k'\}$ of points are obtained with $x_k' p_i^n = y_k' p_i^n$, $(k = 1, 2, \cdots)$ and $\lim_{k\to\infty} x_k' = \lim_{k\to\infty} y_k' = f(p_i^n)$.

Since $p_i^n f(p_i^n) < \min(ab, \epsilon)$ and $\lim_{k\to\infty} x_k' y_k' = 0$, then for k sufficiently large, $p_i^n x_k' = p_i^n y_k' < \epsilon$, but $x_k' y_k' < p_i^n x_k'$, in contradiction to the preceding lemma.

Case 2. The point $f(p_i^n)$ coincides with b. Since $p_i^n b \leqq p_i^n p_{i+j}^n < \lambda(n) < ab$, then $p_i^n \neq a$ and so $i \neq 0$. Denote by $f(b)$ a foot of b on $A(a, p_i^n)$. Then $bf(b) \leqq bp_i^n \leqq p_i^n p_{i+j}^n < \lambda(n) < ab$, and consequently $f(b) \neq a$. The same argument used in Case 1 may be applied to show the existence of two points x, y of $A(a, b)$, on opposite sides of $f(b)$, such that $bx = by < \epsilon$, and $xy < bx$, in contradiction to Lemma 4.1.

If P is any finite subset of a metric space, let R denote any non-reflexive, symmetric, binary relation defined in P, with respect to which each two elements of P are comparable. The subset P is R-connected if and only if P contains for each pair s, t of distinct points, a subset $p_1, p_2, \cdots, p_n$ with $s = p_1$, $t = p_n$ and p_i, p_{i+1} in the relation R for every $i = 1, 2, \cdots, n-1$. The R-content $\mu_R(P)$ is defined to be the sum of the distances of all point-pairs of P that are in the relation R, while the *content* $\mu(P)$ of P is the smallest of the numbers $\mu_R(P)$ for all of the (obviously finite) relations R with respect to which P is R-connected.

COROLLARY. *If A is a metric arc with finite Menger curvature at each point, then for almost all positive integers n, the length of L_n equals $\mu[L_n]$, where $\sum_{i=1}^{n} p_{i-1}^n p_i^n$ defines the length of the n-lattice L_n.*

It is known that $L[P] = \mu(P)$ for any homogeneous chain P of a metric arc, and for almost all natural numbers n, each n-lattice of arc A is, as proved above, a homogeneous chain.†

5. Uniqueness of lattices. Let A be an arc of a metric space M, and f a homeomorphism from $I = [0, 1]$ to A. If $x, y \in A$, we write $x \prec y$ if and only if $f^{-1}(x) < f^{-1}(y)$. Obviously, A is ordered with respect to the relation $\prec$.

If $p \in A$ and $\delta > 0$, $U^R(p; \delta)$ denotes the set $[q \mid q \in U(p; \delta) \cdot A \ \& \ p \prec q]$; similarly $U^L(p; \delta) = [q \mid q \in U(p; \delta) \cdot A \ \& \ q \prec p]$. The arc A is called *metrically monotone* if and only if $x, y, z \in A$, $x \prec y \prec z$ imply $xz \geqq \max(xy, yz)$.

LEMMA 5.1. *If A is a metric arc with finite Menger curvature at each point, then positive numbers δ_R, δ_L exist such that for every point p of A,* (i) *$x, y \in U^R(p; \delta_R)$, $x \prec y$ imply $px < py$, and* (ii) *$x, y \in U^L(p; \delta_L)$, $x \prec y$ imply $px > py$.*

Proof. To prove (i), assume the contrary. Then a sequence $\{\delta_n\}$ of positive numbers exists with $\lim_{n\to\infty} \delta_n = 0$, such that for each positive integer n, A con-

† The proof that $L[P] = \mu[P]$ for any homogeneous chain P is due to Menger. See Distance Geometry, pp. 64–65.

tains a point p_n and $x_n, y_n \in U^R(p_n; \delta_n)$ such that $x_n \prec y_n$, but $p_n x_n \geqq p_n y_n$. Since A is closed and compact, a point p of A and a sub-sequence $\{p_{n_k}\}$ of $\{p_n\}$ exist with $p = \lim_{k\to\infty} p_{n_k}$.

Now corresponding to each positive integer k, the set $U^R(p_{n_k}; \delta_k)$ contains a point y'_{n_k} such that $p_{n_k} y'_{n_k} = p_{n_k} y_{n_k}$, where $p_{n_k} \prec y'_{n_k} \prec x_{n_k}$ in case $p_{n_k} x_{n_k} > p_{n_k} y_{n_k}$, and $y'_{n_k} = x_{n_k}$ if $p_{n_k} x_{n_k} = p_{n_k} y_{n_k}$. The function $\theta(z) = p_{n_k} z$ is continuous on $A(y'_{n_k}, y_{n_k})$, the sub-arc of A with endpoints y'_{n_k}, y_{n_k}, which consequently contains a point m_k that maximizes $\theta(z)$. Consider the infinite sequence $\{m'_k\}$, where $m'_k = m_k$ in case m_k is an interior point of $A(y'_{n_k}, y_{n_k})$. In the contrary case, let f_k denote a foot of p_{n_k} on $A(y'_{n_k}, y_{n_k})$. If f_k is an interior point of that sub-arc, put $m'_k = f_k$, while if f_k is an endpoint select any point in the interior of the sub-arc as m'_k.

Clearly $\lim_{k\to\infty} y'_{n_k} = \lim_{k\to\infty} y_{n_k} = \lim_{k\to\infty} m'_k = p$, and so, for k_0 sufficiently large, pm'_{k_0} is less than the ϵ of Lemma 4.1. It is now easily seen that $A \cdot U(p; pm_{k_0})$ contains points p_{n_k}, r, t (for k sufficiently large) such that $r \in A(y'_{n_k}, m'_k)$, $t \in A(m'_k, y_{n_k})$, $r \neq t$, $p_{n_k} r = p_{n_k} t > rt$, in contradiction to Lemma 4.1.

The argument for (ii) is entirely analogous.

COROLLARY. *If A is a metric arc with finite Menger curvature at each point, then A is the sum of a finite number of non-overlapping, metrically monotone arcs.*

THEOREM 5.1. *If a metric arc A has finite Menger curvature at each point, then there exists a positive integer N such that for $n > N$ the arc A contains exactly one n-lattice.*†

Proof. The existence of at least one n-lattice in A for every natural number n was established by Schoenberg, as remarked above. To prove uniqueness, select N so that $n > N$ implies that $\lambda(n)$ is less than either of the numbers δ_R, δ_L of the preceding lemma, and suppose there are two n-lattices $L^1_n = (p_0, p_1, \cdots, p_n)$, $L^2_n = (q_0, q_1, \cdots, q_n)$ of A with $\lambda_1(n) \leqq \lambda_2(n)$, and $n > N$.

If $\lambda_1(n) < \lambda_2(n)$ then since $p_0 = q_0 = a$ and $\lambda_1(n) < \delta_R$, $\lambda_2(n) < \delta_R$, the points p_1, q_1 belong to $U^R(a; \delta_R)$ and hence $ap_1 < aq_1$ implies $p_1 \prec q_1$. Now $p_1 q_2 > \lambda_2(n)$, for in the contrary case, p_1, $q_1 \in U^L(q_2; \delta_L)$ and $p_1 \prec q_1$ gives $p_1 q_2 > q_2 q_1 = \lambda_2(n)$. It follows that $p_2 \prec q_2$, for, by what has just been established, $p_2 \neq q_2$ while if $q_2 \prec p_2$ then the interior of $A(p_1, q_2)$ contains a point p'_2 such that $p_1 p'_2 = p_1 p_2 = \lambda_1(n)$. Since $p'_2 \prec q_2$ and $q_2 \prec p_2$, then $p'_2 \prec p_2$. But $p_2, p'_2 \in U^R(p_1; \delta_R)$ and so $p_1 p'_2 < p_1 p_2$.

Making the inductive hypothesis that $p_i \prec q_i$ (anchored for $i = 1$), the above procedure shows how $p_n \prec q_n$ follows from $p_{n-1} \prec q_{n-1}$. But since $p_n = q_n = b$, this is impossible.

Finally, use of Lemma 5.1 shows that $p_i = q_i$ $(i = 0, 1, \cdots, n)$ in case $\lambda_1(n) = \lambda_2(n)$, and the theorem is proved.

---

† In a forthcoming paper, the writers obtain the conclusions of Theorems 4.1, 5.1 with considerably weaker hypotheses.

Remark. The restriction of finite Menger curvature at each point enters in the application of Lemma 4.1 to the proof of Lemma 5.1. That the condition is not necessary for uniqueness of n-lattices, for all sufficiently large values of n, is shown by considering the metric arc obtained from the unit segment $[0, 1]$ by defining the distance of each two points as the positive square root of their euclidean distance. Clearly an $(n+1)$-tuple of this arc is an n-lattice if and only if it is an n-lattice of the segment, and hence n-lattices are unique for every n. On the other hand, the metric arc (which, as is well-known, is congruent with an arc of Hilbert space) has *infinite* Menger curvature at each point.

6. Arcs with zero Menger curvature. *Historical remarks.* In this section we are concerned with metric arcs A such that $p \in A$ implies $K(p) = 0$. Calling such arcs "un-curved," the question immediately arises: *Are un-curved arcs straight*; that is, if A is a metric arc, and $p \in A$ implies $K(p) = 0$, is A a metric segment (*i.e.*, congruent with a line segment whose length equals the distance of the end-points of A)? An easy example shows that this question must be answered in the negative. For the circle S_1 (curve), with the distance of two distinct points defined as the length of the shorter arc joining them, is a metric space. Any arc A of S_1, *greater than a semi-circle*, is clearly not a metric segment, though the arc has zero Menger curvature at each point. This example is instructive from another aspect, for not only is arc A un-curved, it is even *locally straight* (that is, A is a *geodesic* arc) but fails to be a metric segment. Menger has given an example of an un-curved metric arc that is not locally straight, and Schoenberg has shown how a wide class of un-curved, not straight, metric arcs may be obtained.

In what metric spaces are un-curved arcs metric segments? The first results in this direction were obtained by Menger, who showed that in any euclidean space (indeed, in any space with each four points congruently imbeddable in euclidean three-space) metric segments are characterized among all arcs by the property of being un-curved.† The highly ingenious proof establishing this important result is based upon two lemmas. One of these lemmas is the n-lattice theorem, referred to in Section 2; the other one (of an *ad hoc* nature) we are not concerned with here.

It was soon pointed out that the argument given for the n-lattice theorem was not valid, and a proof for euclidean arcs was sketched by Alt and Beer. When Schoenberg established the theorem in its present general environment, he showed also that the form in which the second of Menger's two lemmas entered into the proof of the characterization theorem was valid in spaces more general than those possessing the euclidean four-point property; namely the so-called metric ptolemaic spaces.‡ These are metric spaces with the property that for every quadruple p, q, r, s, of points, the three products $pq \cdot rs$, $pr \cdot qs$, $ps \cdot qr$ of

† Distance Geometry, pp. 80–84.

‡ Distance Geometry, pp. 79–80.

"opposite" distances satisfy the triangle inequality. It was concluded, therefore, that in every metric ptolemaic space, un-curved arcs are straight.

Concerning this conclusion, two comments seem pertinent. (1). Statements occur in Menger's proof of the theorem itself (apart from the two lemmas) that are readily acceptable in the case he treated (euclidean space), but which require substantiation when the space is made more general. An instance of this is the expression for arc length that we have obtained in Theorem 3.1. (2). It is tacitly assumed in Menger's proof that an infinite sequence of n-lattices L_n of A exists such that, for each lattice of the sequence, not every three consecutive points have zero curvature. Though no mention of it is made, this assumption need not be verified; but it is easy to show that a ***euclidean*** arc for which it is not valid is a segment. It turns out (as we shall show) that metric ptolemaic arcs for which this assumption is not valid are also metric segments. It happens that metric betweenness possesses a transitive property in metric ptolemaic spaces that it does not have in merely metric spaces. By virtue of this transitivity, the desired property of arcs will be established in the next section.

The problem of characterizing metric segments as un-curved arcs of a restricted class of metric spaces was studied also by Haantjes.†

In the second section of Finsler's dissertation a definition of curvature is formulated that, as Haantjes observed, is directly applicable to rectifiable arcs of any metric space, though it is not, perhaps, so purely a metrization of the notion as is Menger's definition. If $p \in A(a, b)$, a rectifiable metric arc, the Finsler-Haantjes curvature $k(p)$ of $A(a, b)$ at p is given by

$$k(p) = \left\{ 24 \lim_{q,r \to p} (L[A(q, r)] - qr)/L^3[A(q, r)] \right\}^{1/2},$$

$q, r \in A(a, b)$ where $L[A(q, r)]$ denotes the length of that sub-arc of $A(a, b)$ with endpoints q, r.

This notion of curvature is defined only for rectifiable arcs, while no such restriction is explicit in the definition of Menger curvature (the notion of which is meaningful at each accumulaton point of any metric space, provided the limit involved exists). But the greater generality of the latter definition is more apparent than real, for Pauc has shown that if p is a point of any metric continuum and $K(p)$ exists, then in a neighborhood of p the continuum is a rectifiable arc.‡ It follows that if $K(p)$ exists at each point p of a metric arc $A(a, b)$, then the arc is rectifiable and the Finsler-Haantjes curvature notion is applicable.

Haantjes proved that for metric arcs the existence of both $K(p)$ and $k(p)$ implies their equality; that if $K(p)$ exists so does $k(p)$, but not conversely; while for euclidean arcs the two notions are equivalent. On the basis of Pauc's theorem and the second of the above results, he proved that a metric, ptolemaic arc with

† Distance Geometry, pp. 76–78.

‡ C. Pauc, Courbure dans les espaces métriques, Rend. Accad. Naz. die Lincei, vol. 24, 1936, pp. 109–115. The theorem is announced without proof. In a forthcoming paper by the writers, the theorem is seen to follow as a corollary of a much more general theorem.

$K(p) \equiv 0$ is a metric segment. His proof of this fundamental theorem is quite rigorous, but in view of the number of preliminary theorems upon which it is based (some of which are not easily established) it is far from being direct or complete in itself.

7. A new approach to the fundamental theorem. In view of the remarks made in the preceding section, it seems desirable to obtain a proof of the theorem characterizing segments in metric ptolemaic spaces as un-curved arcs that is direct, complete in itself, and free from lacunae. We indicate now how this may be done; the details will appear in another publication.

An arc is called a *geodesic* provided each of its points is contained in a subarc that is a metric segment. The following lemma, needed in our approach, is of intrinsic interest and fills the gap (mentioned in Section 6) in the Menger-Schoenberg procedure.

LEMMA 7.1. *In a metric ptolemaic space, each geodesic arc is a metric segment.*

Proof. If $A(a, b)$ is a geodesic arc, it follows from its definition and a standard compactness argument that a positive number δ exists such that for each point p of the arc, $U(p; \delta) \cdot A(a, b)$ is linear (that is, congruently imbeddable in a straight line). Let N be a positive integer such that $n > N$ implies that for every n-lattice L_n of $A(a, b)$, $\lambda(n) < \delta$. Then each consecutive triple of points in every n-lattice L_n, $(n > N)$, is linear and, since $p_i^n p_j^n = \lambda(n)$ for $|i-j| = 1$, it follows that p_{i+1}^n is metrically between p_i^n and p_{i+2}^n for $i = 0, 1, \cdots, n-2$ (that is, $p_i^n \neq p_{i+1}^n \neq p_{i+2}^n$ and $p_i^n p_{i+1}^n + p_{i+1}^n p_{i+2}^n = p_i^n p_{i+2}^n$). Symbolize this betweenness relation by $p_i^n p_{i+1}^n p_{i+2}^n$.

Consider the quadruple p_0^n, p_1^n, p_2^n, p_3^n. Since the space is ptolemaic,

$$p_0^n p_1^n \cdot p_2^n p_3^n + p_0^n p_3^n \cdot p_1^n p_2^n \geqq p_0^n p_2^n \cdot p_1^n p_3^n.$$

Now $p_0^n p_1^n = p_1^n p_2^n = p_2^n p_3^n = \lambda(n)$, and $p_0^n p_2^n = p_1^n p_3^n = 2\lambda(n)$ (since $p_0^n p_1^n p_2^n$ and $p_1^n p_2^n p_3^n$ hold), and so the above inequality yields $p_0^n p_3^n \geqq 3\lambda(n)$. But since the space is metric, $p_0^n p_3^n \leqq p_0^n p_1^n + p_1^n p_2^n + p_2^n p_3^n = 3\lambda(n)$; that is, $p_0^n p_3^n = 3\lambda(n)$.

Making the inductive assumption that $p_0^n p_k^n = k \cdot \lambda(n)$ for $k = 2, 3, \cdots, n-1$, and applying the above procedure to the quadruple p_0^n, p_{n-2}^n, p_{n-1}^n, p_n^n, we obtain $p_0^n p_n^n = n \cdot \lambda(n)$. Hence for $n > N$,

$$p_0^n p_n^n / [n \cdot \lambda(n)] = ab / [n \cdot \lambda(n)] = 1.$$

Taking the limit as $n \to \infty$, use of Theorem 3.1 gives $L[A(a, b)] = ab$; that is, the length of $A(a, b)$ is the distance of its end-points, and hence $A(a, b)$ is a metric segment.

The argument used in the proof of Lemma 7.1 shows that if $A(a, b)$ is a metric ptolemaic arc and there is an infinite sequence of n-lattices of $A(a, b)$ such that each consecutive triple of points of each n-lattice of the sequence has zero curvature, then $A(a, b)$ is a metric segment.

Remark. To prove that a metric ptolemaic arc with everywhere vanishing Menger curvature is a metric segment, it suffices to show that it is a geodesic; moreover, according to the Corollary of Lemma 5.1, we may restrict our attention to metrically monotone arcs.

LEMMA 7.2. *If A is any metric arc with zero curvature at each point, then to each $\epsilon>0$ there corresponds a positive integer N such that for $n>N$ the curvature of each consecutive triple of points of every n-lattice of A is less than ϵ.*

Proof. Since $K(p)=0$, $p\in A$, there corresponds to each p of A a positive number $\sigma(p)$ such that every three points of $A\cdot U(p;\sigma(p))$ has curvature less than ϵ, and the usual compactness argument shows that $\sigma(p)>\sigma>0$ for every p of A. It suffices, then, to take N such that $n>N$ implies $\lambda(n)<\sigma$, for then p_i^n, p_{i+1}^n, $p_{i+2}^n\in A\cdot U(p_{i+1}^n;\sigma)$, and consequently $K(p_i^n, p_{i+1}^n, p_{i+2}^n)<\epsilon$ for $n>N$.

Now let $A(a, b)$ be a monotone arc of a metric ptolemaic space, and for a given positive (arbitrarily small) ϵ, let N be a positive integer such that (1) for $n>N$ *each* consecutive triple p_i^n, p_{i+1}^n, p_{i+2}^n of each n-lattice $L_n=(a=p_0^n, p_1^n, \cdots, p_n^n=b)$ has curvature less than ϵ, and (2) $\lambda(n)<\epsilon/2^m$, (m arbitrarily large). It follows that

$$1 \geqq p_i^n p_{i+2}^n/[2\lambda(n)] \geqq 1-\epsilon^2\lambda^2(n).$$

Application of the ptolemaic inequality to an arbitrary quadruple p_i^n, p_{i+1}^n, p_{i+2}^n, p_{i+3}^n of consecutive points of L_n gives

$$(*)\qquad 1 \geqq p_i^n p_{i+3}^n/[3\lambda(n)] \geqq 1-(4/3)\lambda^2(n)[\epsilon_{i+1}+\epsilon_{i+2}-\lambda^2(n)\epsilon_{i+1}\epsilon_{i+2}],$$

where $\epsilon_{i+j}=\epsilon_{i+j}(n)\geqq 0$, $\lim_{n\to\infty}\epsilon_{i+j}=0$, $(j=1, 2)$.

The proof is completed by extending (*) inductively to obtain

$$\lim_{n\to\infty} p_0^n p_n^n/n\cdot\lambda(n) = 1.$$

THE APPROXIMATION OF LOGARITHMIC NUMBERS

H. T. DAVIS, Northwestern University

1. Introduction. The coefficients of the powers of x in the expansion of $x/\log(1+x)$ are called *logarithmic numbers*. Let us denote these by L_n, so that

$$\frac{x}{\log(1+x)} = 1 + L_1x + L_2x^2 + L_3x^3 + \cdots + L_nx^n + \cdots, \tag{1}$$

where $L_1=\frac{1}{2}$, $L_2=-1/12$, $L_3=1/24$, $L_4=-19/720$, *etc.*

These numbers to $n=20$ have been computed by A. N. Lowan and H. E. Salzer,* who terminated their table with the value

$$L_{20} = -\frac{12365\ 72232\ 34699\ 80029}{48\ 17145\ 97618\ 97472\ 00000} = -0.00256\ 7022545\cdots.$$

The size of the numerator and denominator of this fraction and the obvious, though not insuperable, difficulties of extending the computation to higher values of n suggest the problem of searching for an asymptotic formula for logarithmic numbers. Such a formula was published in 1924 by J. F. Steffensen,† who gave as an asymptotic approximation the following:

$$L_n \sim \frac{(-1)^{n+1}}{n\log^2 n} = S_n. \tag{2}$$

Unfortunately the ratio L_n/S_n converges very slowly toward 1 as one sees from the values $L_{20}/S_{20}=0.461\cdots$, $L_{100}/S_{100}=0.545\cdots$. In fact, for the ratio L_n/S_n to vary from 1 by as little as 0.01, n would have to exceed 29×10^{12}.

The principal object of this paper is to improve the formula of Steffensen and to suggest other methods by means of which more satisfactory approximations can be made to L_n for large values of n. Incidentally, these numbers will be generalized by means of what we shall call the *logarithmic number function*, from which a companion set of numbers corresponding to $n=m+\frac{1}{2}$ will be defined, where m is an integer. These numbers will be found to have asymptotic properties similar to those of L_n.

2. Properties of logarithmic numbers. Logarithmic numbers find their principal use in integration formulas. Thus, if we denote by x_m the quantity $x_0+m\omega$ and by f_m the value $f(x_m)$, where $f(x)$ is a function integrable over the interval (x_0, x_p), then the Gregory-Newton formula for finite integration is written as follows:

* Tables of coefficients in numerical interpolation formulae, Journal of Math. and Physics, vol. 22, 1943, pp. 49–50.

† On Laplace's and Gauss' summation-formulas, Skandinavisk Aktuariettidskrift, 1924, pp. 2–4. See also Interpolation, Baltimore, 1927, pp. 106–107.

$$(3)\qquad \frac{1}{\omega}\int_{x_0}^{x_p} f(s)ds = \sum_{m=0}^{p} f_m - L_1(f_0 + f_p) + L_2(\Delta f_{p-1} - \Delta f_0) - L_3(\Delta^2 f_{p-2} + \Delta^2 f_0) + L_4(\Delta^3 f_{p-3} - \Delta^3 f_0) + \cdots.$$

Logarithmic numbers are also special values of Bernoulli polynomials of nth order and nth degree, that is,

$$(4)\qquad L_n = \frac{B_n^{(n)}(1)}{n!},$$

where $B_m^{(n)}(x)$ is the coefficient of $t^m/m!$ in the expansion of the function $t^n e^{xt}/(e^t - 1)^n$.

Observing the expansion

$$\frac{x}{(1+x)\log(1+x)} = \sum_{m=0}^{\infty} \frac{x^m B_m^{(m)}(0)}{m!},$$

we have from comparison with (1) the equation

$$(5)\qquad L_n = [nB_{n-1}^{(n-1)}(0) + B_n^{(n)}(0)]/n!.$$

From the identity

$$B_n^{(n)}(x) = \int_x^{x+1} (s-1)(s-2)\cdots(s-n)ds,$$

we then obtain, by means of (5), the following essential formula:

$$(6)\qquad L_n = \int_0^1 \frac{s(s-1)\cdots(s-n+1)}{n!}\,ds.$$

In order to attain a more satisfactory form for L_n, we write

$$s(s-1)\cdots(s-n+1) = \Gamma(s+1)/\Gamma(s-n+1), \qquad n! = \Gamma(n+1).$$

By means of these identities we can express (6) as follows:

$$(7)\qquad L_n = \int_0^1 \frac{\Gamma(s+1)}{\Gamma(n+1)\Gamma(s-n+1)}\,ds.$$

If, furthermore, we observe the identity

$$\Gamma(x)\Gamma(1-x) = \pi/\sin \pi x,$$

and use it to write

$$\Gamma(s-n+1)=\Gamma[1-(n-s)]=\frac{\pi}{\sin\pi(n-s)\Gamma(n-s)}$$

(8)

$$=\frac{\pi}{(-1)^{n+1}\sin\pi s\,\Gamma(n-s)},$$

then the integral (7) can be reduced to the following attractive form

$$L_n=\frac{(-1)^{n+1}}{\pi}\int_0^1\frac{\Gamma(s+1)\sin\pi s\,\Gamma(n-s)}{\Gamma(n+1)}\,ds. \tag{9}$$

If one makes the observation that the integrand of (9) is actually the function $\sin\pi sB(s+1, n-s)$, where $B(p, q)$ is the Beta function, then one can express L_n as the double integral

$$L_n=\frac{(-1)^{n+1}}{\pi}\int_0^1\int_0^1\sin\pi s\,t^s(1-t)^{n-s-1}dtds.$$

The representation of L_n given by (9) is readily adapted to the evaluation of L_n by means of finite integration. Within the limits of integration $\Gamma(s+1)$ varies from a minimum value

$$\gamma_0=\Gamma_n(1+s_0)=0.88560\ 31943,\ \text{at}\ s_0=0.46163\ 21450, \tag{10}$$

to a maximum value of 1 at $s=0$ and $s=1$. The ratio $\Gamma(n-s)/\Gamma(n+1)$ varies from a maximum of $1/n$ at $s=0$ to a minimum of $1/n(n-1)$ at $s=1$.

The rapidity of the convergence of the integral is shown from the fact that the values of L_{20} to 9 significant figures and of L_{100} to 7 significant figures were computed by means of formula (3) in which only eleven values of the integrand were used, namely, those over the range $s=0(.1)1$. For future reference the latter value to the indicated approximation is recorded as follows:

$$L_{100}=-\,0.000297\ 4763\cdots. \tag{11}$$

3. Asymptotic formulas for logarithmic numbers. We now have in (9) a formula readily adapted to achieve the desired asymptotic representation of L_n. We first replace $\Gamma(n-s)$ and $\Gamma(n+1)$ by their Stirling approximations, namely,

$$\Gamma(x)\sim x^{x-1}(2\pi x)^{1/2}e^{-x}\left[1+\frac{1}{12x}+\cdots\right].$$

Making the proper substitutions and simplifying the calculation, we obtain

$$\frac{\Gamma(n-s)}{\Gamma(n+1)}\sim\frac{e^{-as}}{n}\left(1-\frac{s}{n}\right)^{-s+1/2}\sim\frac{e^{-as}}{n},$$

where $a=\log n$.

When this value is substituted in (9), we obtain the following asymptotic expression for L_n:

$$L_n \sim \frac{(-1)^{n+1}}{n\pi} \int_0^1 \Gamma(s+1) \sin \pi s \, e^{-as} ds. \tag{12}$$

Since the integrand is positive we can now apply the law of the mean and thus obtain

$$L_n \sim \frac{(-1)^{n+1}}{n\pi} \Gamma(\xi+1) \int_0^1 \sin \pi s \, e^{-as} ds, \qquad 0 \leqq \xi \leqq 1, \tag{13}$$

and hence,

$$L_n \sim \frac{(-1)^{n+1}\Gamma(\xi+1)}{n(\log^2 n + \pi^2)} = T_n \Gamma(\xi+1). \tag{14}$$

Using the extreme values of $\Gamma(\xi+1)$, namely 1 and γ_0 given by (10), we obtain the following inequalities for $n=20$ and $n=100$:

$$0.002350 < L_{20} < 0.002653, \qquad 0.0002850 < L_{100} < 0.0003218. \tag{15}$$

Since $L_{20}/T_{20}=0.967$ and $L_{100}/T_{100}=0.924$, one conjectures that the limiting value of L_n/T_n will be close to γ_0. This has not been proved, however.

Although the asymptotic formula given by (14) is a reasonably satisfactory one, and does indeed establish the ultimate correctness of Steffensen's approximation given in Section 1, it would be useful to achieve a sharper one if possible. An obvious suggestion is to replace $\Gamma(s+1)$ in (12) by its power series representation. Unfortunately, the proximity of the singular point at $s=-1$ makes the series converge very slowly in the interval (0, 1) and actually diverge at $s=1$. These difficulties can be largely surmounted, however, if a polynomial approximation to $\Gamma(s+1)$ is made by means of Lagrange's interpolation formula, or some other similar device.

Actually, it was found desirable to compute the cubic polynomial which has the value 1 at $s=0$ and $s=1$, which has the value γ_0 at $s=s_0$, and the derivative of which is zero at $s=s_0$. This polynomial is readily found to be

$$y = 1 - a_1 s + a_2 s^2 - a_3 s^3,$$

where $a_1=0.52590\ 56748$, $a_2=0.66802\ 76014$, $a_3=0.14212\ 19266$.

When this function is introduced into (12), we then have

$$L_n \sim \frac{(-1)^{n+1}}{n\pi} \left[I_0(a) - a_1 I_1(a) + a_2 I_2(a) - a_3 I_3(a) \right], \tag{16}$$

where we introduce the function

$$I_p(a) = \int_0^1 s^p \sin \pi s \, e^{-as} ds.$$

It will be useful also to introduce the companion function

$$J_p(a) = \int_0^1 s^p \cos \pi s \, e^{-as} ds.$$

These functions are readily shown to satisfy the following relationships:

$$I_p(a) = \frac{\pi e^{-a}}{a^2+\pi^2} + \frac{p}{a^2+\pi^2}\left[aI_{p-1}(a) + \pi J_{p-1}(a)\right],$$
$$J_p(a) = \frac{\pi e^{-a}}{a^2+\pi^2} + \frac{p}{a^2+\pi^2}\left[aJ_{p-1}(a) - \pi I_{p-1}(a)\right]. \tag{17}$$

They also satisfy separately the following iterative equations:

$$I_{p+2}(a) = \frac{\pi e^{-a}}{a^2+\pi^2} + \frac{2a(p+2)}{a^2+\pi^2} I_{p+1}(a) - \frac{(p+1)(p+2)}{a^2+\pi^2} I_p(a).$$
$$J_{p+2}(a) = \frac{a-p-2}{a^2+\pi^2} e^{-a} + \frac{2a(p+2)}{a^2+\pi^2} J_{p+1}(a) - \frac{(p+1)(p+2)}{a^2+\pi^2} J_p(a). \tag{18}$$

Beginning with the initial values,

$$I_0 = \frac{\pi(e^{-a}+1)}{a^2+\pi^2}, \qquad J_0 = \frac{a(e^{-a}+1)}{a^2+\pi^2},$$
$$I_1 = \frac{\pi e^{-a}}{a^2+\pi^2} + \frac{1}{a^2+\pi^2}(aI_0 + \pi J_0), \quad J_1 = \frac{ae^{-a}}{a^2+\pi^2} + \frac{1}{a^2+\pi^2}(aJ_0 - \pi I_0), \tag{19}$$

the functions of higher subscript are readily computed from (18). Equations (17) are useful in checking the computations.

By means of these formulas the efficiency of the approximation function (16) was checked numerically for $n=20$ and $n=100$. The following values for I_p $(p=0, 1, 2, 3)$ were obtained:

$n = 20$	$n = 100$
$I_0 = 0.17505\ 14455$	$I_0 = 0.10210\ 08618$
$I_1 = 0.01390\ 15517$	$I_1 = 0.03127\ 05043$
$I_2 = 0.03045\ 03062$	$I_2 = 0.01297\ 53716$
$I_3 = 0.01700\ 51089$	$I_3 = 0.00651\ 01003$

When these values are substituted in (16), the approximations $L_{20} \sim .00254 \cdots$ (with an error of 3 in the last place) and $L_{100} \sim .0002973$ (with an error of 1 in the last place) are obtained.

4. Inequalities satisfied by logarithmic numbers. In order to establish an inequality between successive logarithmic numbers, we shall write $n+1$ for n in formula (9) and replace $(-1)^{n+1}$ by $\cos(n+1)\pi$. We then have

$$L_{n+1} = -\frac{\cos(n+1)\pi}{\pi\Gamma(n+2)}\int_0^1 \Gamma(s+1)\sin \pi s\,\Gamma(n+1-s)ds,$$

$$= -\frac{\cos(n+1)\pi}{\pi(n+1)\Gamma(n+1)}\int_0^1 \Gamma(s+1)\sin \pi s\,(n-s)\,\Gamma(n-s)ds,$$

$$= -\frac{\cos(n+1)\pi}{\pi\Gamma(n+1)}\int_0^1 \Gamma(s+1)\sin \pi s\,\Gamma(n-s)ds \tag{20}$$

$$+\frac{\cos(n+1)\pi}{\pi(n+1)\Gamma(n+1)}\int_0^1 \Gamma(s+2)\sin \pi s\,\Gamma(n-s)ds,$$

$$= -L_n + \frac{\cos(n+1)\pi}{\pi(n+1)\Gamma(n+1)}\int_0^1 \frac{\Gamma(s+2)}{\Gamma(s+1)}\Gamma(s+1)\sin \pi s\Gamma(n-s)ds.$$

But by the theorem of mean value, the last integral is equal to

$$\frac{\Gamma(\xi+2)}{\Gamma(\xi-1)}\,\frac{L_n}{n+1} = \frac{\xi+1}{n+1}L_n, \qquad 0 \leqq \xi \leqq 1,$$

so that, from (20) we get

$$L_{n+1} = -L_n\left[1-\frac{\xi+1}{n+1}\right], \qquad 0 \leqq \xi \leqq 1. \tag{21}$$

Letting ξ have successively the values 0 and 1, we obtain the following inequality:

$$\frac{n}{n+1}\,|L_n| \geqq |L_{n+1}| \geqq \frac{n-1}{n+1}\,|L_n|. \tag{22}$$

By an argument which does not differ essentially from that just given, we obtain the following expression for L_{n+p}:

$$L_{n+p} = (-1)^p L_n\left[1-\frac{\xi_1+1}{n+p}\right]\left[1-\frac{\xi_2+1}{n+p-1}\right]\cdots\left[1-\frac{\xi_p+1}{n+1}\right], \qquad 0 \leqq \xi_i \leqq 1.$$

If all the ξ_i are set equal to 1, we get the inequality

$$|L_{n+p}| \geqq \frac{n(n-1)}{(n+p)(n+p-1)} \geqq |L_n|;$$

and if all the ξ_i are set equal to 0, we have

$$|L_{n+p}| \leqq \frac{n}{n+p}\,|L_n|;$$

that is to say, we have established the following inequality, of which (22) is a special case:

$$\frac{n}{n+p}\,|L_n| \geqq |L_{n+p}| \geqq \frac{n(n-1)}{(n+p)(n+p-1)}\,|L_n|. \tag{23}$$

5. The logarithmic number function. If we return to equation (7) we see that the integrand is defined over a continuous range of values of n. This suggests the possibility of defining a logarithmic number function of the continuous variable n, which would, in particular, reduce to L_n when n is an integer. Denoting this function by $L(n)$, we thus write

$$L(n) = \int_0^1 \frac{\Gamma(s+1)}{\Gamma(n+1)\Gamma(s-n+1)}\,ds, \tag{24}$$

or, making use of (8),

$$\begin{aligned} L(n) &= \frac{1}{\pi\Gamma(n+1)}\int_0^1 \sin\pi(n-s)\,\Gamma(s+1)\Gamma(n-s)ds, \\ &= \pi n P(n) - \cos\pi n\, Q(n), \end{aligned} \tag{25}$$

where we abbreviate

$$P(n) = K(n)\int_0^1 \cos\pi s\,\Gamma(s+1)\Gamma(n-s)ds,$$

$$Q(n) = K(n)\int_0^1 \sin\pi s\,\Gamma(s+1)\Gamma(n-s)ds, \qquad K(n) = 1/[\pi\,\Gamma(n+1)].$$

We first observe the special values: $L(-n) = 0$, $L(0) = 1$, and $L(n) = (-1)^{n+1}Q(n) = L_n$, when n is a positive integer.

If n is replaced by $m+\frac{1}{2}$, where m is an integer, we obtain *half-logarithmic numbers*, which we denote by $L_{m+1/2}$, that is,

$$L(m+\tfrac{1}{2}) = (-1)^m P(m+\tfrac{1}{2}) = L_{m+1/2}. \tag{26}$$

These numbers can be computed from (24) by finite integration. Beginning with $m=-1$, we find the following seven-figure approximation of the first four numbers as follows:

$$L_{-1/2} = 0.5097422, \quad L_{1/2} = 0.9854485, \quad L_{3/2} = 0.03483709, \quad L_{5/2} = 0.007700113.$$

From a consideration of (24) we readily prove that $L(n)$ has one zero in each of the intervals $p<n<p+1$ $(p=1, 2, 3, \cdots)$.

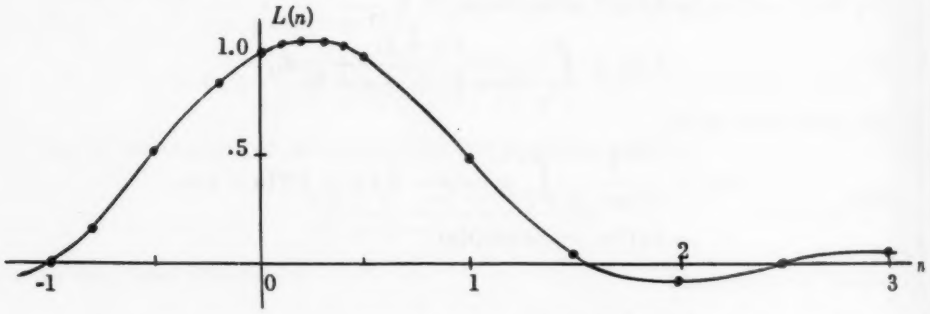
We also see that $L(n)$ has a maximum value between $n=0$ and $n=1$. This is readily shown from the derivative of $L(n)$, that is,

$$\frac{dL}{dn} = L'(n) = \int_0^1 \frac{\Gamma(s+1)}{\Gamma(n+1)\Gamma(s-n+1)}\,[\Psi(s-n+1) - \Psi(n+1)]ds, \tag{27}$$

which is obtained by differentiating (24) and introducing the psi-function defined by the ratio: $\Psi(x)=\Gamma'(x)/\Gamma(x)$.

Since $\Psi(s+1)-\Psi(0)$ is positive and $\Psi(s)-\Psi(2)$ is negative in the interval $0\leqq s\leqq 1$, the derivative changes from a positive to a negative value between $n=0$ and $n=1$. Hence a maximum value exists in the interval. This maximum is found to be in the neighborhood of $s=0.2$ and is approximately 1.068.

The graphical representation of $L(n)$ between $n=-1$ and $n=3$ is shown in the accompanying figure. The points on the curve were obtained by the finite integration of (24).



An asymptotic expansion for $L(n)$ for positive values of n is obtained if we replace $\Gamma(n-s)/\Gamma(n+1)$ in (25) by its asymptotic value, namely, e^{-as}/n, where $a=\log n$. We thus obtain

$$(28)\qquad L(n)\sim\frac{\sin n\pi}{n\pi}\int_0^1\cos\pi s\,\Gamma(s+1)\,e^{-as}ds-\frac{\cos n\pi}{n\pi}\int_0^1\sin\pi s\,\Gamma(s+1)\,e^{-as}ds.$$

If in this expression we replace $\Gamma(s+1)$ by $1+A_1s+A_2s^2+A_3s^3+\cdots$, we obtain the further approximation,

$$(29)\qquad \begin{aligned} L(n)\sim{}&\frac{\sin n\pi}{n\pi}(J_0+A_1J_1+A_2J_2+A_3J_3+\cdots)\\ &-\frac{\cos n\pi}{n\pi}(I_0+A_1I_1+A_2I_2+A_3I_3+\cdots), \end{aligned}$$

where I_p and J_p are the functions defined in Section 3.

If in (29) n is replaced by $m+\frac{1}{2}$, where m is an integer, the coefficient of the second series vanishes and we have an asymptotic formula for the half-logarithmic numbers. Neglecting all but the first term, we obtain

$$(30)\qquad L_{m+1/2}\sim\frac{(-1)^m}{(m+1/2)\pi}J_0\sim(-1)^m\frac{a}{(m+1/2)\pi(a^2+\pi^2)},$$

where $a=\log(m+\frac{1}{2})$.

MATHEMATICAL INDUCTION IN SETS

W. L. DUREN, JR., University of Virginia

Recent years have seen an increased use of a principle of proof which is analogous to ordinary mathematical induction over the natural numbers. This principle, known as the Hausdorff Maximal Principle or Zorn's Lemma, was first recognized to be equivalent to transfinite induction and the axiom of choice and so it was used in abstract algebra and general set theory [1], [2]. However, the same style of proof by a form of mathematical induction over the linearly ordered set of real numbers in an interval $[a, b]$ yields simple and elegant proofs in the theory of the calculus, as pointed out by L. R. Ford [3]. Thus the principle of mathematical induction unifies a diverse set of methods of proof at many levels of mathematics.

Let us first examine a palpably false theorem about natural numbers: *The set of all natural numbers is a finite set.* Attempting the proof by ordinary mathematical induction, we easily show that the set $\{1\}$ is finite and that the cardinal set $\{1, 2, \cdots, n\}$ is finite if the cardinal set $\{1, 2, \cdots, n-1\}$ is finite. Hence by the principle of mathematical induction, for every n the cardinal set $C_n = \{1, \cdots, n\}$ is finite. These cardinal sets C_n which have been proved finite form a chain $\mathfrak{C}$ of sets totally ordered by containment. That is, if C_n and C_m are two sets in the chain, either $C_n \subset C_m$ or $C_m \subset C_n$. Moreover, the chain $\mathfrak{C}$ is a *maximal* totally ordered chain of finite cardinal sets. That is, if another chain of finite cardinal sets exists one of whose members is contained in a member of $\mathfrak{C}$, then every member of this new chain is a member of $\mathfrak{C}$. Now to establish the conclusion of the theorem we must show that the union of all the sets in the chain $\mathfrak{C}$ also has the property of being a finite set. This certainly does not follow and, in fact, it is not true; so the "proof" fails.

In the proofs by general mathematical induction one forms a maximal chain of sets each of which has some property P and the issue which characteristically arises is whether the union of the chain has the property P. In proofs of theorems of the calculus the form of induction which is applicable is induction over the linearly ordered (but not well ordered) set of real numbers in an interval, and our characteristic issue is often settled by means of boundedness of the interval and the Dedekind property of the real numbers. Let us examine a proof of the Heine-Borel theorem which is a prototype for proofs of this sort. [4], [5].

The Heine-Borel theorem for the real numbers states that if a family $\mathfrak{F}$ of open intervals covers a closed and bounded interval $[a, b]$ of real numbers, then a finite subfamily of the covering family also covers the interval $[a, b]$. (Here a family of open intervals is said to *cover* the set $[a, b]$ if and only if every point x in the interval is contained in some open interval of the family.)

We make the proof by mathematical induction over the linearly ordered set of real numbers in the interval. First, by hypothesis, the initial point a is contained in some open interval of the family. Hence some non-empty open

interval $I_1=\{a\leqq x<c_1\}$ is covered by a finite subcollection (namely one) of the sets of the covering family. Now we form a maximal chain $\mathfrak{C}$ of sets $I=\{a\leqq x<c\}$ with $I_1\subset I\subset[a, b]$, each having the property that it is covered by a finite subcollection of open intervals in $\mathfrak{F}$. Does the union of this chain which is an interval $I_0=\{a\leqq x<c_0\}$ have a finite covering? We answer this question by the aid of the Dedekind property. The number c_0 is the least upper bound of numbers c such that $\{a\leqq x<c\}$ has a finite covering. Since the interval $[a, b]$ is closed the number c_0 is in $[a, b]$. Hence c_0 is in some open interval J of the family $\mathfrak{F}$. J intersects some one of the intervals $I=\{a\leqq x<c\}$ which has a finite covering and thus $I\cup J$ has a finite covering. Moreover, $I\cup J$ contains I_0 so that the union of the chain *does* have the property of being finitely covered. Now if $c_0=b$ the proof is complete and if $c_0<b$ interval $I\cup J$ properly contains I_0 so that the chain $\mathfrak{C}$ cannot be a maximal one. Hence the interval $[a, b]$ has a finite covering.

The above wording is not quite the simplest statement of the argument but it emphasizes the issue about the maximal chain of sets with the required property. It is worthy of note that the method of proof does not make use of the separability of the numbers, that is, the fact that the countable set of rational numbers is dense in the interval $[a, b]$. In fact, the method of proof yields actual generalizations of the Heine-Borel theorem for real numbers. In particular, we find that neither the statement nor the proof of the Heine-Borel theorem needs modification to apply it to the interval $[-\infty, +\infty]$ in the extended real numbers. For this interval is "bounded" in the sense required by the proof by mathematical induction.

It is also instructive to attempt to replace the covering family $\mathfrak{F}$ of open intervals, $\{x_1<x<x_2\}$ by a covering family $\mathfrak{F}'$ of half-open intervals $\{x_1\leqq x<x_2\}$. A statement analogous to the Heine-Borel theorem can be made and the proof proceeds as before until we reach the statement: "J intersects some one of the intervals $I=\{a\leqq x<c\}$, which has a finite covering." This is false for the half-open covering intervals. So we cannot establish that the union of the maximal chain of sets with finite covering also has a finite covering. In fact, the analogue of the Heine-Borel theorem for half-open intervals is false.

Let us apply the induction style of argument to obtain a direct proof of the theorem:

If f is a real valued function on the closed interval $[a, b]$ and if the derivative $f'=0$ at every point of $[a, b]$ then f is constant.

Proof: For every $\epsilon>0$ there is an interval $I_1=\{a\leqq x<c_1\}$ such that in I_1, $f(x)-f(a)<\epsilon(x-a)$. This follows from the definition of the derivative and the fact that $f'(a)=0$. As before we form a maximal chain $\mathfrak{C}$ of intervals $I=\{a\leqq x<c\}$ with $I_1\subset I\subset[a, b]$ in each of which $f(x)-f(a)<\epsilon(x-a)$. The union of the chain $\mathfrak{C}$ is an interval $I_0=\{a\leqq x<c_0\}$. Taking the least upper bound c_0 of I_0 we see that $f(c_0)-f(a)\leqq\epsilon(c_0-a)$. Since c_0 is in $[a, b]$ and hence $f'(c_0)=0$ we also have, for all x in some open interval J containing c_0, $f(x)-f(c_0)<\epsilon(x-c_0)$. Adding the last two inequalities we find that, for all x in J, $f(x)-f(a)<\epsilon(x-a)$. Moreover J

intersects one of the intervals I in $\mathfrak{C}$ and hence the same inequality holds in $I \cup J$ which properly contains I_0, so the union of the chain *does* have the property. If $c_0 < b$ the chain $\mathfrak{C}$ cannot be maximal. Hence $c_0 = b$ and for all x in $[a, b]$ we have $f(x) - f(a) \leqq \epsilon(b-a)$. Analogously for all x in $[a, b]$ we have $f(x) - f(a) \geqq \epsilon(b-a)$. Since ϵ is arbitrary this implies that $f(x) \equiv f(a)$.

The method can be generalized according to an observation of L. R. Ford [3], if we define an *extendable property* of open sets.

DEFINITION. *A property P of open sets is an extendable property if for every pair A, B of open sets each with property P, and having non-void intersection, their union $A \cup B$ also has property P.*

Now by applying the method of proof by induction over the linearly ordered set of real numbers in an interval $[a, b]$, we readily prove the following theorem due to L. R. Ford.

FORD'S THEOREM. *If P is an extendable property of open sets and if every point x in the bounded and closed interval $[a, b]$ is contained in an open set having property P, then the interval $[a, b]$ has property P.*

More generally, defining a set A to be *compact* if every covering of A by a collection of open sets contains a finite subcollection which covers A, we can prove by generalized mathematical induction the theorem: *If a compact set A is covered by a family of open sets each having the extendable property P, then A has the property P.* However, this generalization does not contain all instances of extension of an extendable property of sets by mathematical induction.

In applying mathematical induction over the linearly ordered interval of real numbers $[a, b]$ both the assertion of the existence of a maximal chain $\mathfrak{C}$ of sets having the extendable property P and the technique for showing that the union of the maximal chain has the property P depend upon the Dedekind property, the existence of the least upper bound. In ordinary (finite) induction over the natural numbers the existence of a maximal chain $\mathfrak{C}$ having property P depends upon the fact that the natural numbers are *well ordered*, which implies that if the theorem is false for some n, then there exists a first number n_1 for which the theorem is false. However, the fact that the union of any maximal chain of subsets of a finite set is a member of the chain comes from the finiteness itself. In the case of transfinite induction the existence of a maximal chain of sets having property P follows from the well-ordering but the technique of extension past the limit ordinals is similar to the calculus proofs using the Dedekind property. Finally, in general set theory [1], [2], the existence of the maximal chain of sets having the property P is a separate logico-mathematical principle, and the question whether the union of the maximal chain has property P, and the question of the existence of a proper extension of it having property P, are still ones for which there are no general techniques for answering.

Let us return in conclusion to one more example from the calculus. This is the theorem that a continuous function f on a bounded and closed real interval $[a, b]$ is uniformly continuous in $[a, b]$. To prove this theorem we must first

show that ordinary continuity implies local uniform continuity, which is a property of sets. That is, we first show that for every point x_0 in the interval and every $\epsilon>0$ there exists a $\delta>0$ such that if x_1 and x_2 are points of $[a, b]$ in a δ-neighborhood of x_0, then $|f(x_1)-f(x_2)|<\epsilon$. This is proved by observing that ordinary continuity at x_0 implies that for every $\epsilon/2>0$ there exists a $\delta/2>0$ such that if x_1 is in $[a, b]$ and $|x_1-x_0|<\delta/2$ then $|f(x_1)-f(x_0)|<\epsilon/2$. Similarly for x_2. Then the local uniform continuity at x_0 follows from combining these inequalities. But the uniform continuity of f is an extendable property of open intervals, hence by Ford's theorem it is true in the interval $[a, b]$.

References

1. E. R. Stabler, This MONTHLY, vol. 51, 1944, p. 129.
2. O. Frink, This MONTHLY, vol. 61, 1954, p. 223.
3. L. R. Ford, Interval additive propositions (abstract), Bull. Am. Math. Soc. vol. 62, 1956, p. 158; This MONTHLY, vol. 64, 1957, p. 107.
4. Veblen and Lennes. Infinitesimal Analysis. N. Y. 1907, p. 34.
5. T. H. Hildebrandt. Bull. Am. Math. Soc. vol. 32, 1926, p. 423.

CALCULATION OF MOMENTS FOR A CANTOR-VITALI FUNCTION

G. C. EVANS, University of California

1. The "middle-third" function. Our task is to find moments of the form $M_n=\int_0^1 x^n p(x)dx$ and the related moments $m_n=\int_0^1 x^n dp(x)$ for functions $p(x)$ which have certain properties of symmetry. The most familiar of these is the one which is based on the Cantor "middle-third" set, which is obtained from the interval $0\leqq x\leqq 1$ by removing first the open middle third $1/3<x<2/3$, then, from the remaining closed intervals, their open middle thirds $1/9<x<2/9$ and $7/9<x<8/9$, and so on, from remaining closed intervals, open middle thirds indefinitely. The union of this denumerable collection of open intervals constitutes an open set E, of which the remaining complement F is the typical Cantor *discontinuum*. Between any two points of F there lie intervals of E; F is a perfect nowhere dense set. A simple calculation shows that the total of the lengths of the intervals of E is the value 1 itself.

The corresponding function $p(x)$ is defined first for points of E: as 1/2 on $1/3<x<2/3$; as 1/4 on $1/9<x<2/9$, as 3/4 on $7/9<x<8/9$; as 1/8 on $1/27<x<2/27$, as 3/8 on $7/27<x<8/27$, $\cdots$. Thus on E, $p(x)$ is a nondecreasing

function of x, constant on each interval of E (see Figure). Next we write $p(0) = 0$, $p(1) = 1$, and to the end points of each interval of E assign to $p(x)$ the value it has for the interval itself. In order to complete the definition on remaining points of F we note that for the nondecreasing function so far defined the values $y = p(x)$ are dense in $0 \leqq y \leqq 1$; in fact, at a stage k of setting up the intervals of E, y has been given values, $m/2^k$, $m = 1, 2, \cdots, 2^k - 1$. Hence at a point of F

there is one and only one value y which will make the function $p(x)$ continuous at the point. We assign that value. It follows that $p(x)$ is monotonic-increasing on $0 \leqq x \leqq 1$.*

* If the reader has not run across the fact, it will interest him to prove that, in the scale of 3 the x-coordinates of the right hand end points of intervals of E are arbitrary decimal fractions composed of a finite number of digits 2 and 0. The corresponding y values, in the scale of 2, however, are given by the same sequences except that digits 2 are replaced by 1. For example, when $x = .2022$, $y = .1011$.

The points of F, in the scale of 3, are given by all decimal fractions, finite or infinite, which can be written as sequences of digits 2 and 0; from this it follows by a conventional argument that F is a nondenumerable infinite set of points. The points $.0222 \cdots = .1 = 1/3$, $.020222 \cdots = .021 = 7/27$ are in F as left end points of intervals of E, and $020202 \cdots = 1/4$ is in F but is not an end point. In fact the end points of intervals are only denumerable.

Functions of the kind considered in this article are merely very special examples of the type of nondecreasing continuous function, called by Vitali an *elementary discard*, for which is assigned a dense set of constant values of y on a sequence of intervals which exhausts in measure the fundamental interval. In terms of elementary discards Vitali analyzes the singular part of a continuous function of bounded variation (Rendiconti del Circolo Matematico di Palermo, vol. 46, 1922, p. 388).

The Stieltjes integrals $\int_0^1 f(x)dp(x)$, $\int_0^1 p(x)df(x)$, if $f(x)$ is continuous, exist as limits of Riemann sums, and

$$\int_0^1 f(x)dp(x) = f(x)p(x)\Big]_0^1 - \int_0^1 p(x)df(x). \tag{1}$$

From (1) we have

$$\int_0^1 x^{n+1}dp(x) = x^{n+1}p(x)\Big]_0^1 - \int_0^1 p(x)(n+1)x^n dx.$$

Hence

$$m_{n+1} = 1 - (n+1)M_n. \tag{2}$$

For this "middle-third" function $p(x)$, as a special case, we shall derive the formula

$$2M_n = \frac{1}{n+1} + \frac{1}{3^{n+1}-1}\{(2+M)^n - M_n\},$$

expressing in symbolic form a binomial development in which M^k stands for M_k, and M^0 for M_0 instead of 1.*

2. The class of functions $p(x)$. As a slight generalization, we consider the functions $p(x)$ in which the number 3 of §1 is replaced by an a, rational or irrational, with $a>1$.

From $0 \leqq x \leqq 1$ we take the middle open interval of length $d_1 = 1/a$, and from the two remaining closed intervals, each of length $f_1 = (1/2)(1-1/a)$, the middle open intervals of length $d_2 = \{(a-1)/2a\}(1/a)$, and so on successively, from each of the closed intervals of length f_k the middle open portions of length $d_{k+1} = f_k/a$ thus, at stage k,

$$d_k = \left(\frac{a-1}{2a}\right)^{k-1}\frac{1}{a}, \qquad f_k = \left(\frac{a-1}{2a}\right)^k.$$

Since there are 2^{k-1} open intervals taken in passing from stage $k-1$ to stage k, the total of lengths d_k, $k=1, 2, 3, \cdots$, is

$$\sum_1^\infty 2^{k-1}d_k = 1.$$

The union of the open intervals of lengths $d_1, d_2, \cdots$ constitute the set E, and the intersection of the closed intervals (adding the points $x=0$, $x=1$) the set F, which is the complement of E.

The function $p(x)$ is then defined in complete analogy with the function of §1: first on the points of E, as 1/2 on the central interval of stage 1, as 1/4 and

* This MONTHLY (abstract) vol. 58, 1951, p. 593.

3/4 on those of stage 2, and so on, and then extended to the points of F, to obtain finally a monotonic increasing function, continuous for $0 \leqq x \leqq 1$.

We find that $p(x)$ has the following properties:

$$p(x) = \frac{1}{2} + p\left(x - \frac{a+1}{2a}\right), \qquad \frac{a+1}{2a} \leqq x \leqq 1, \tag{3}$$

$$p(x) = 2p\left(\frac{a-1}{2a}\, x\right), \qquad 0 \leqq x \leqq 1, \tag{4}$$

$$p(x) + p(1-x) = 1, \qquad 0 \leqq x \leqq 1. \tag{5}$$

Of these, (5) is immediate on account of the symmetry of the construction, and (3) is a statement that starting from the right-hand end point $x = 1/2 + 1/2a$, of the interval of stage 1 of E, the values assigned to $p(x)$ to the right are $1/2$ plus the values assigned to the right of $x = 0$.

The relation (4) also is made clear by the method of construction.* The restriction of the function $p(x)$ to the interval $0 \leqq x \leqq 1/2 - 1/2a$ is a scale model of $p(x)$ on $0 \leqq x \leqq 1$, with the x-axis in the scale $(a-1)/2a$: 1 and the y-axis in scale 1:2. Relation (4) is allied to the determination of the metrical dimension of F.†

3. Calculation of M_n, m_n. We use (3) and (4) to calculate M_n. We have

$$M_n = \int_0^1 x^n p(x)\,dx = \int_0^{(a-1)/2a} x^n p(x)\,dx + \int_{(a+1)/2a}^1 x^n p(x)\,dx + \int_{(a-1)/2a}^{(a+1)/2a} \frac{1}{2}\, x^n dx.$$

where by (3)

$$\int_{(a+1)/2a}^1 x^n p(x)\,dx = \int_{(a+1)/2a}^1 \frac{1}{2}\, x^n dx + \int_{(a+1)/2a}^1 x^n p\left(x - \frac{a+1}{2a}\right) dx.$$

By means of changes of variable, $t = x - (a+1)/2a$ in the last integral above, and $t = x$ in the first integral in the expression for M_n, we have

* The following sentence is a substitution, suggested by the referee, for a longer proof.

† Physical dimension is an analysis of the problem of change of units. Let new units of length, time and mass have the measures l, k, and n respectively in terms of given units. If the measure of a quantity in terms of the new units is $l^{-r}k^{-s}n^{-t}$ times its measure in terms of the old ones, the quantity has dimensions r, s, t respectively in length, mass and time. A generalization of this idea in terms of measure or mass of sets leads to the concept of metric, or Hausdorff, dimension. One sees that since "half" the set F lies on a portion $(a-1)/2a$ of the original interval, the metric dimension D of F is log 2 to the base $2a/(a-1)$. For $a = 3$, $D = \log_3 2$; for $a = 2$, $D = \frac{1}{2}$.

The easiest way to make the procedure legitimate is to define a mass measure of F. Let there be a uniform distribution of unit mass on $0 \leqq x \leqq 1$. At the first stage move this mass from the central interval of length $1/a$, half to be distributed uniformly on each of the adjacent intervals, constituting F_1, and continue this process indefinitely, in the stage m moving from a central interval of F_{m-1} half the mass uniformly to each of the contiguous intervals of F_m. Thus in the limiting situation the mass on the set $x_1 \leqq x \leqq x_2$ is $p(x_2) - p(x_1)$. Then (4) may be used.

In more general cases one may use methods associated with Vitali's theorem. For a discussion of metric dimension see Hurewicz and Wallman, Dimension Theory, Princeton, 1948, Chapter VII.

$$M_n = \int_0^{(a-1)/2a} \left\{ t^n + \left(t + \frac{a+1}{2a} \right)^n \right\} p(t)\,dt + \frac{1}{2} \int_{(a-1)/2a}^{1} x^n dx$$

$$= \frac{a-1}{2a} \int_0^1 \left(\frac{a-1}{2a} \right)^n \left\{ x^n + \left(x + \frac{a+1}{a-1} \right)^n \right\} p\left(\frac{a-1}{2a}\, x \right) dx$$

$$+ \frac{1}{2(n+1)} \left\{ 1 - \left(\frac{a-1}{2a} \right)^{n+1} \right\}.$$

By (4),

$$p\left(\frac{a-1}{2a}\, x \right) = \frac{1}{2}\, p(x).$$

With this substitution, and a rearrangement which moves the terms in x^n of the expansion in the right member into the left hand member, we obtain the equation

$$\left\{ \left(\frac{2a}{a-1} \right)^{n+1} - 1 \right\} M_n$$

$$= \frac{1}{2(n+1)} \left\{ \left(\frac{2a}{a-1} \right)^{n+1} - 1 \right\} + \frac{1}{2} \sum_{k=0}^{n-1} \binom{n}{k} \left(\frac{a+1}{a-1} \right)^{n-k} \int_0^1 x^k p(x)\,dx.$$

Thus finally,

$$2M_n = \frac{1}{n+1} + \frac{1}{\left(\frac{2a}{a-1} \right)^{n+1} - 1} \sum_{k=0}^{n-1} \binom{n}{k} \left(\frac{a+1}{a-1} \right)^{n-k} M_k. \tag{6}$$

This formula may be written in the form

$$2M_n = \frac{1}{n+1} + \frac{1}{\left(\frac{2a}{a-1} \right)^{n+1} + 1} \left\{ \left(M + \frac{a+1}{a+1} \right)^n - M_n \right\}, \tag{7}$$

where in the binomial development M^k stands for M_k and M^0 for M_0 instead of 1, as in the special case $a=3$ at the end of §1. The values m_n are then given by (2).

In particular, we find

$$M_0 = \frac{1}{2}, \qquad M_1 = \frac{1}{2}\,\frac{2a-1}{3a-1}, \qquad M_2 = \frac{9a-5}{12\,(3a-1)},$$

and from the fact that $p(0)=0$, $p(1)=1$,

$$m_0 = 1.$$

4. The symmetry relation (5). In the above calculation it has not been necessary to use (5). By means of it we may write $M_n = \int_0^1 x^n p(x)dx$, after change of variable $t = 1 - x$, in the form

$$M_n = \int_0^1 (1-t)^n p(1-t)dt = \int_0^1 (1-t)^n \{1 - p(t)\}dt$$

$$= \int_0^1 (1-t)^n dt - \int_0^1 (1-t)^n p(t)dt.$$

In this way we have

$$M_n = \frac{1}{n+1} + \sum_{k=0}^{n} (-1)^{k+1} \binom{n}{k} M_k,$$

(8)

$$M_n(1 + (-1)^n) = \frac{1}{n+1} + \sum_{k=0}^{n-1} (-1)^{k+1} \binom{n}{k} M_k,$$

valid whenever (5) is valid. From (8), which does not involve a, the moments M_n of even subscripts are determined successively in terms of those of odd subscript. It may be used as a check on (6). Also, since the M_n itself is eliminated when n is odd, there are two expressions involving the same moments.

The three relations (3), (4), (5), with the requirement of continuity do not determine the function $p(x)$ uniquely, because there remains the possibility of assigning on the middle interval, of length $1/a$, other values than $y = 1/2$. Evidently (3) and (5) do not imply (4). On the other hand, although (3) is convenient to use, it is a consequence of (4) and (5).

In fact, by application in succession of (4), (5), (4), and (5), with $(a+1)/2a \leqq x \leqq 1$, we have

$$\frac{1}{2} + p\left(x - \frac{a+1}{2a}\right) = \frac{1}{2} + \frac{1}{2}\, p\left(\frac{2a}{a-1}\, x - \frac{a+1}{a-1}\right),$$

$$0 \leqq x - \frac{a+1}{2a} \leqq \frac{a-1}{2a}, \quad 0 \leqq \frac{2a}{a-1}\, x - \frac{a+1}{a-1} \leqq 1,$$

$$= 1 - \frac{1}{2}\, p\left(\frac{2a}{a-1}(1-x)\right), \qquad 0 \leqq \frac{2a}{a-1}(1-x) \leqq 1,$$

$$= 1 - p(1-x), \qquad 0 \leqq 1 - x \leqq \frac{a-1}{2a},$$

$$= p(x).$$

The ranges of the variables are noted in order to make clear the geometry of the transformations.

SOLUTION OF A RANKING PROBLEM FROM BINARY COMPARISONS

L. R. FORD, JR., RAND Corporation

1. Introduction. Occasionally one is faced with the problem of ranking a collection of objects on the basis of a number of binary comparisons, where these have not been, nor cannot be, chosen in a manner desirable to the experimenter. The standard procedures (*e.g.*, p. 217 of [1]) seem to require usually that the number of comparisons between any given pair be equal to the number between any other pair; for one reason or another it may be quite difficult in practice to achieve this equality for a suitably large sample size. For example, if one were attempting to rank various makes of automobiles by obtaining paired comparisons only from persons who have actually owned both makes one sees easily that it might be extremely difficult to obtain a large number of comparisons for such pairs as Cadillac *versus* Crosley. On the other hand, Chevrolet *versus* Ford comparisons would be available in large numbers.

The method proposed in this paper does not require any specific number of comparisons between pairs, and for this reason may have some application in contexts similar to the above. Another potential application is in ranking players of a two-person game on the basis of their records against each other. In collegiate football, for example, where matches are not scheduled in a manner suited to the experimenter, this technique could be used.

We assume as given a matrix $A=(a_{ij})$, where a_{ij} represents the number of times object i has been preferred to object j ($a_{ii}=0$). The approach we shall use is that of maximum likelihood. We associate with the ith object a weight w_i. These weights will be interpreted as odds, in the sense that the probability of i being preferred to j in a future comparison will be taken to be $w_i/(w_i+w_j)$.

With these probabilities we may compute the *a priori* probability of obtaining precisely the matrix of results which we in fact did obtain (*i.e.*, the matrix A), under the assumption that each comparison that takes place between i and j is independent of all other comparisons, and is drawn randomly from a binomial distribution with $p=w_i/(w_i+w_j)$ and with the number of ij comparisons being required to be equal to $a_{ij}+a_{ji}$.

We shall then attempt to assign w_i that are positive and that maximize this probability. The function to be maximized is:

$$\Pr(A \mid w_1, \cdots, w_N) = \prod_{i<j} \binom{a_{ij}+a_{ji}}{a_{ij}} \left(\frac{w_i}{w_i+w_j}\right)^{a_{ij}} \left(\frac{w_j}{w_i+w_j}\right)^{a_{ji}}.$$

We shall actually solve the following equivalent problem: In the set $\{w_i>0;\ \sum w_i=1\}$, find values w_i that maximize

$$F_A(w) = \prod_{i<j} \left(\frac{w_i}{w_i+w_j}\right)^{a_{ij}} \left(\frac{w_j}{w_i+w_j}\right)^{a_{ji}}.$$

We shall make the following assumption about the matrix A:

ASSUMPTION. *In every possible partition of the objects into two nonempty subsets, some object in the second set has been preferred at least once to some object in the first set.*

This assumption is a requirement that one might reasonably expect. For if it is violated then there exist two subsets S_1 and S_2 with either no interset comparisons, or with all interset comparisons favoring S_1. In the former case one cannot expect to rank any object in S_1 against any object in S_2, since there is no basis for comparison. In the latter case the maximizing weights assigned to S_2 must turn out to be zero, since otherwise we could increase our function $F_A(w)$ by multiplying the S_1 weights by some common factor >1 and multiplying the S_2 weights by some common factor <1. These factors may be chosen to preserve a sum of 1. The factors of $F_A(w)$ involving ij both in S_1 or both in S_2 remain unchanged by this operation. However, the factors involving i in S_1 and j in S_2 increase, thereby increasing $F_A(w)$.

But if the S_2 weights must all be zero, we cannot hope to compare individual members of S_2 by this process.

The partition assumption is sufficient to establish the existence of a unique set of w_i in the region $\{w_i<0;\ \sum w_i=1\}$ which maximize $F_A(w)$. An equivalent form of the partition assumption which we shall find useful is the following: For each ordered pair ij there exists a sequence of indices $i_0, i_1, \cdots, i_n$ with $i_0=i$, $i_n=j$ and with $a_{i_k i_{k+1}}>0$ for $k=0, 1, \cdots, n-1$.

2. Existence and uniqueness of the solution. For the remainder of the paper, we shall assume that our result matrix $A=(a_{ij})$ satisfies the assumption of the preceding section.

We readily see that $F_A(w)$ is positive in the region $\{w_i>0;\ \sum w_i=1\}$. Thus to establish the existence of a maximum in the region it will suffice to show that defining $F_A(w)=0$ for w on the boundary gives a continuous extension of $F_A(w)$ to the closed region $\{w_i\geqq 0;\ \sum w_i=1\}$.

If $w^0=(w_1^0, \cdots, w_N^0)$ is on the boundary, then there is an index i such that $w_i^0=0$ and a j such that $w_j^0>0$. By our assumption on A we can find a sequence of indices $i_0, i_1, \cdots, i_n$ with $i_0=i$, $i_n=j$, and $a_{i_k i_{k+1}}>0$, $k=0, \cdots, n-1$. Clearly there are adjacent indices in the sequence, say m and k, with $w_m^0=0$, $w_k^0>0$, and $a_{mk}>0$. We may now write

$$F_A(w)=\left(\frac{w_m}{w_m+w_k}\right)^{a_{mk}}\cdot\phi(w),$$

where $0<\phi(w)\leqq 1$ in the interior. Hence $\lim_{w\to w^0} F_A(w)=0$ and continuity of the extension follows; hence also the existence of a maximum in the interior follows.

Taking logarithms and partial derivatives leads to the following system of equations which w must satisfy at a maximum point:

$$W_i = \sum_{j=1}^{N} \frac{A_{ij} w_i}{w_i + w_j}, \qquad i = 1, 2, \cdots, N, \tag{1}$$

where $W_i = \sum_j a_{ij}$ and $A_{ij} = a_{ij} + a_{ji}$. Let $\bar{w}$ be any interior point satisfying equations (1). Define a new matrix $C = (c_{ij})$ by $c_{ij} = A_{ij}\bar{w}_i/(\bar{w}_i + \bar{w}_j)$. Using the matrix C we may compute Υ_i and C_{ij} (corresponding to W_i and A_{ij}) and we see that

$$\Upsilon_i = \sum_j c_{ij} = \sum_j \frac{A_{ij}\bar{w}_i}{\bar{w}_i + \bar{w}_j} = W_i, \tag{2}$$

$$C_{ij} = c_{ij} + c_{ji} = A_{ij}\bar{w}_i/(\bar{w}_i + \bar{w}_j) + A_{ji}\bar{w}_j/(\bar{w}_j + \bar{w}_i) = A_{ij}.$$

However, note also that we may write, using relations (2),

$$F_A(w) \equiv \frac{\prod_i w_i^{W_i}}{\prod_{i<j} (w_i + w_j)^{A_{ij}}} \equiv \frac{\prod_i w_i^{\Upsilon_i}}{\prod_{i<j} (w_i + w_j)^{C_{ij}}} \equiv F_C(w), \tag{3}$$

so that $F_A(w) \equiv F_C(w)$ over the entire region. We shall now need the following result.

LEMMA. *If l_{ij} and l_{ji} are both positive, then*

$$\left(\frac{w_i}{w_i + w_j}\right)^{l_{ij}} \left(\frac{w_j}{w_i + w_j}\right)^{l_{ji}}$$

has a maximum (unique up to a proportionality factor) at the values $w_i = K \cdot l_{ij}$, $w_j = K \cdot l_{ji}$.

Our assumption on the A matrix now implies that $F_C(w)$ has a unique maximum at $\bar{w}$, as follows. Suppose we assign the weight $\bar{w}_1$ to the first object. For the jth object we have a sequence $i_0, i_1, \cdots, i_n$, with $i_0 = 1$, $i_n = j$, and $a_{i_k i_{k+1}} > 0$, $k = 0, 1, \cdots, n-1$. Hence, since $\bar{w}$ is an interior point, $c_{i_k i_{k+1}}$ and $c_{i_{k+1} i_k}$ are both positive. We proceed inductively along this sequence, assigning weights *via* the above lemma to maximize the factor of the form

$$\left(\frac{w_{i_k}}{w_{i_k} + w_{i_{k+1}}}\right)^{c_{i_k i_{k+1}}} \left(\frac{w_{i_{k+1}}}{w_{i_k} + w_{i_{k+1}}}\right)^{c_{i_{k+1} i_k}}$$

at each stage. We see that having assigned $w_{i_k} = \bar{w}_{i_k}$ we must assign $w_{i_{k+1}} = \bar{w}_{i_{k+1}}$ in order to maximize. Hence w_j must be taken to be $\bar{w}_j$, and thus uniqueness of the maximum of $F_C(w)$ at $\bar{w}$ follows.

3. The iteration and proof of convergence.* Since there appears to be no simple direct way of solving equations (1), we resort to the following iterative scheme:

* Standard conditions for the convergence of a scheme of this sort (*e.g.*, pp. 206–210 of [2]) are not satisfied in our case.

$$(4)\qquad \left.\begin{aligned} w_i^{n+1} &= \frac{W_i}{\sum_j \frac{A_{ij}}{w_i^n + w_j^n}} \\ w_j^{n+1} &= w_j^n, \quad j \neq i. \end{aligned}\right\} \quad i = 1, 2, \cdots, N \text{ cyclically.}$$

Clearly, if we start with $w_i^0 > 0$ we retain this property in subsequent iterations. Furthermore, comparison of (4) with (1) indicates that if we can demonstrate convergence to an interior point (*i.e.*, a point in $\{w_i > 0;\ \sum w_i = 1\}$), the limit must be the desired solution. (Properly speaking, the limiting values will not satisfy $\sum w_i = 1$. Equations (1) are homogeneous of degree zero in the w_i, and we shall actually obtain convergence of the ratios of the w_i. We could normalize at each stage of the iteration; this was deemed unnecessary, since the w_i are important only up to a proportionality factor.)

Assume that we are about to apply (4) to compute a new w_i. The index i will remain fixed in what follows. Define

$$(5)\qquad \delta_i(w) = \frac{\partial \ln F_A(w)}{\partial w_i} = \frac{W_i}{w_i} - \sum_k \frac{A_{ik}}{w_i + w_k},$$

$$(6)\qquad D_i(w) = w_i \cdot \delta_i(w) = W_i - \sum_k A_{ik} \frac{w_i}{w_i + w_k}.$$

From (6) we observe that $D_i(w)$ is monotone in w_i for $0 < w_i < \infty$. Hence if we can show that $\delta_i(w^n)$ and $\delta_i(w^{n+1})$ have the same sign, we may assert that this sign prevails along the entire line segment between w^n and w^{n+1}. For this purpose notice from (5) and (4) that

$$\begin{aligned} \delta_i(w^{n+1}) &= \frac{W_i}{w_i^{n+1}} - \sum_k \frac{A_{ik}}{w_i^{n+1} + w_k^{n+1}} \\ &= \sum_k \frac{A_{ik}}{w_i^n + w_k^n} - \sum_k \frac{A_{ik}}{\dfrac{W_i}{\sum_j \dfrac{A_{ij}}{w_i^n + w_j^n}} + w_k^n} \\ &= \sum_k A_{ik} \left\{ \frac{1}{w_i^n + w_k^n} - \frac{\sum_j \dfrac{A_{ij}}{w_i^n + w_j^n}}{W_i + w_k^n \sum_j \dfrac{A_{ij}}{w_i^n + w_j^n}} \right\} \\ &= \delta_i(w^n) \left\{ \sum_k \frac{A_{ik} w_i^n}{(w_i^n + w_k^n)\left(W_i + w_k^n \sum_j \dfrac{A_{ij}}{w_i^n + w_j^n}\right)} \right\} \end{aligned}$$

and thus that $\delta_i(w)$ does not change sign along the segment joining w^n to w^{n+1}.

Now we note from (4), (5), and (6) that

$$\delta_i(w^n) = \frac{[w_i^{n+1} - w_i^n]W_i}{w_i^{n+1} \cdot w_i^n} \tag{7}$$

and conclude from this, since our assumption guarantees that $W_i > 0$, that $\delta_i(w^n)$ is positive or negative according as $w_i^{n+1} - w_i^n$ is positive or negative. But now, using the fact that $\delta_i(w)$ is $\partial \ln F_A(w)/\partial w_i$ and does not change sign along the segment from w^n to w^{n+1}, we see that $F_A(w)$ is monotone non-decreasing as w_i moves from w^n to w^{n+1}, and is strictly monotone if $\delta_i(w^n) \neq 0$.

Since $F_A(w)$ is continuous and is equal to zero on the boundary, we see that the successive points arrived at under the iterations are bounded away from the boundary.

The iteration proceeds cyclically. Let w_0 be the initial normalized approximation. Let w_K be the normalized w-vector after K complete cycles from 1 to N. As we have just seen, $F_A(w)$ increases under an i change unless $\delta_i(w) = 0$, in which case none of the quantities involved is altered. Hence, if $F_A(w)$ is unchanged during a complete cycle, $\delta_i(w)$ must be zero for every i. However, $F_A(w_K)$ is monotone in K and bounded; hence $F_A(w_K)$ approaches a limit. Let $\bar{w}$ be any limit point of the w_K sequence.

Since the change in $F_A(w)$ during a complete cycle is a continuous function of w, the change in $F_A(\bar{w})$ must be zero, and thus $\delta_i(\bar{w}) = 0$, for all i. But this is precisely the statement that equations (1) are satisfied at $\bar{w}$. However, $\bar{w}$ was any limit point of w_K, and equations (1) have a unique solution in the interior.

Thus $w_K \to \bar{w}$ and the proof of convergence to the unique interior point satisfying (1) is complete.

4. Some properties and examples. We note the following: If each of the $(N^2 - N)/2$ possible comparisons has been made the same number of times (*i.e.*, $A_{ij} = Q$) then the ordering obtained from the weights is the same as that obtained from comparing the total percentages of favorable comparisons (the win-loss percentages).

For if $W_i > W_j$ then, at a solution,

$$0 < \sum_k \frac{Qw_i}{w_i + w_k} - \sum_k \frac{Qw_j}{w_j + w_k} = Q(w_i - w_j) \sum_k \frac{w_k}{(w_i + w_k)(w_j + w_k)},$$

and $w_i > w_j$ follows. This indicates, for example, that the system of ranking used in major-league baseball is in agreement with that proposed in this paper. In this case it should be noted, however, that the win-loss percentages are not in general a solution of the weight equations, but merely give the same ordering.

In case the A_{ij} are not constant, the example

$$A = \begin{pmatrix} 0 & 15 & 15 & 0 \\ 11 & 0 & 10 & 20 \\ 11 & 10 & 0 & 20 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

shows that the win-loss percentages do not necessarily represent the desired rankings:

Object	Win-Loss %	Weight
1	.576	.736
2	.611	.541
3	.611	.541
4	.047	.025

Here objects 2 and 3 have the highest win-loss percentages as a result of their many wins over the weak object 4; but object 1, which has consistently beaten objects 2 and 3, has the greatest weight.

References

1. Guildford, J. P., Psychometric Methods, McGraw-Hill Book Company, Inc., New York, 1936.

2. Scarborough, J. B., Numerical Mathematical Analysis, Johns Hopkins Press, Baltimore, 1950.

ON THE COMPUTATION OF CERTAIN ROOTS BY THE USE OF THE BINOMIAL SERIES*

JOHN W. GREEN, The University of California

When the author was a student in Professor Ford's calculus class at the Rice Institute, the subject under discussion on one occasion was that of finding roots —in this instance, cube roots—of numbers by using the binomial expansion. The class was invited to propose numbers of which to compute the cube root, and the number 37 was suggested. A student's slide rule gave the approximate value 3 1/3 for the cube root of 37, and this was taken as point of departure for the binomial expansion:

$$\begin{aligned}\sqrt[3]{37} &= \frac{10}{3}\sqrt[3]{37\cdot\left(\frac{3}{10}\right)^3} = \frac{10}{3}\sqrt[3]{\frac{37\cdot 27}{1000}} \\ &= \frac{10}{3}\sqrt[3]{\frac{999}{1000}} = \frac{10}{3}\sqrt[3]{1-\frac{1}{10^3}} \\ &= \frac{10}{3}\left\{1-\frac{1}{3\cdot 10^3}-\frac{1\cdot 2}{3^2\cdot 10^6\cdot 2!}-\frac{1\cdot 2\cdot 5}{3^3\cdot 10^9\cdot 3!}-\cdots\right\}.\end{aligned}$$

The series is a delight to compute and converges very rapidly; the first two terms give

$$\sqrt[3]{37} = 3.33222\cdots,$$

correct to five decimal places.

Presumably, the student's happy choice of the number 37 was a fortuitous one. It appears of some interest, if only pedagogical, to look into the reason why things worked out like they did, and to look for further examples of this sort of phenomenon. "This sort of phenomenon" will be defined as the occurrence of an integer K for which

$$\sqrt[3]{K} = \frac{p}{q}\sqrt[3]{1\pm\frac{1}{10^n}}, \tag{1}$$

p, q, and n being integers.

To simplify the analysis, consider first the analogous problem for the square root:

* The preparation of this paper was sponsored in part by the Office of Naval Research and the Office of Ordnance Research, U. S. Army. Reproduction in whole or in part is permitted for any purpose of the U. S. Government.

$$\sqrt{K} = \frac{p}{q}\sqrt{1 \pm \frac{1}{10^n}}. \tag{2}$$

We may assume without loss of generality that p and q have no other common divisor than 1 ($(p, q) = 1$) and that K is square free. From (2) it follows that

$$q^2K = p^2 \pm \frac{p^2}{10^n}. \tag{3}$$

Case 1. $n = 2m$ is even.

Clearly $p^2/10^n = p^2/10^{2m}$ is an integer, and, since it is the square of a rational number, it is the square r^2 of an integer r:

$$q^2K = p^2 \pm r^2. \tag{4}$$

From (4) one sees that any factor common to q and r would be a factor of p, and hence $(q, r) = 1$. From (3) and $p^2 = r^2 \cdot 10^{2m}$, it follows that

$$q^2K = r^2(10^{2m} \pm 1).$$

Since r^2 divides q^2K and $(K, q) = 1$, r^2 divides K. Since K has no repeated factors except 1, we conclude that $r = 1$ and

$$q^2K = 10^{2m} \pm 1. \tag{5}$$

Case 2. $n = 2m+1$ is odd.

In this case $p^2/10^{2m+1} = S$, an integer. Then $p^2/10^{2m} = 10S = r^2$, as in Case 1. Now the square of an integer r can not terminate in 0 unless r terminates in 0; thus $r = 10t$, and $p^2/10^{2m+1} = 10t^2$. We note that $(q, t) = 1$ and substitute for p^2 in (3). This gives

$$q^2K = 10t^2(10^{2m+1} \pm 1).$$

As in Case 1, $t = 1$ and we obtain

$$q^2K = 10(10^{2m+1} \pm 1). \tag{6}$$

From (5) and (6) it is evident that the phenomenon in which we are interested can occur when and only when $10^n \pm 1$ has a repeated factor q, and therefore to find examples of it, we must find n such that $10^n \pm 1$ contains a repeated factor.

For the cube root case, the conclusion is analogous. The situation described in (1) can occur only if $10^n \pm 1$ has a twice repeated factor. The case mentioned at the beginning of this article is that corresponding to $n = 3$; $10^3 - 1 = 3^3 \cdot 37$.

There appears to be no general method of factoring numbers of the form $10^n \pm 1$. One would not expect there to be, considering the similarity of these numbers to the Mersenne numbers. Therefore, examples of the phenomena we are seeking may be expected to be hard to come by. A considerable amount of attention, however, has been given to their factorization for special values of n,

and a number of tables exist giving their factors for small n. (See Lehmer, [2].) The author had available the table of Cunningham and Woodall, [1] which gives, among other things, all factors of $10^n \pm 1$ for $n \leqq 20$ except the case $10^{19}+1$.

Examination of this table reveals very few examples, and makes the student's choice of 37 appear even more astonishingly fortunate. In fact, for $n \leqq 20$, no other examples at all of triple factors of $10^n \pm 1$ appear, except for the obvious cases $n = 3k$, which again have the factor 27. The case $n = 6$ gives

$$\sqrt[3]{37037} = \frac{100}{3}\sqrt[3]{1 - \frac{1}{10^6}},$$

and this is the smallest number except 37 for which (1) has been found to hold.

The square root problem also has few examples. Of course, for every n, $10^n - 1$ is divisible by 9, so there are a number of more or less equivalent examples:

$$\sqrt{11} = \sqrt{\frac{99}{9}} = \frac{10}{3}\sqrt{1 - \frac{1}{10^2}},$$

$$\sqrt{1110} = \frac{100}{3}\sqrt{1 - \frac{1}{10^3}},$$

etc. The only example of repeated roots other than 3 occurred in the case $n = 11$,

$$10^{11} + 1 = 11^2 \cdot 826446281,$$

which gives the example

$$\sqrt{8264462810} = \frac{10^6}{11}\sqrt{1 + \frac{1}{10^{11}}}.$$

To obtain further examples requires factorization of further large numbers. However, it may be possible to show by general methods of number theory that certain K may *not* be expressed in forms (1) or (2). For example, in the square root case, simple considerations show that any K satisfying (2) must terminate in digits 1 or 9 for the case n even, or in the digits 10 or 90 in the case n odd. It would be interesting to know that those K expressible in the forms (1) or (2) are really as rare as they seem to be.

Bibliography

1. A. J. C. Cunningham and H. J. Woodall, Factorization of $(y^n \pm 1)$, London, 1925.
2. D. H. Lehmer, Guide to Tables in the Theory of Numbers, Washington, D. C., 1941.

AN EXCERPT FROM THE WORKS OF EULER

RUDOLPH E. LANGER, University of Wisconsin

Among the phenomena that dramatically accent the history of mathematics is the abundance of genius that trod the stage of the eighteenth century. The era immediately preceding that had marked the birth of the great key ideas that unlocked the door to the mathematics of modern times. Mathematical analysis had suddenly been endowed with power and dependability beyond all that had previously been suspected. Seemingly limitless fields for development had been revealed all around, and opportunities for triumphs of ingenuity and invention were abundantly uncovered. The stimulus was matched by the response. Genius, ordinarily so rare, cropped up generously. Great masters of mathematical analysis were contemporaries. In many respects the greatest of these was Leonhard Euler.

Euler needed no throne to sit upon, nor any army to command, to leave his mark upon history. In an age which accorded mathematical ability the primacy as a mark of human greatness, he was the pre-eminent mathematician. Researches and books flowed from his pen in incredible volume. He was a superlative expositor, and an openhearted soul who corresponded—in Latin, German, French and Russian—with practically all of his great contemporaries. Adulation was his in overflowing measure, but left no tarnish upon him. A modest and generous personality, he was content to be a burgher, the jolly father of thirteen children. His afflictions, heavy though they were, did not embitter him. To his death, at the age of seventy-seven, the inspired enthusiasm of the explorer stayed with him. He enjoyed the discoveries of others as he did his own.

Euler was born in 1707. He was therefore nine when Leibniz died, and twenty when Newton followed. His early manhood came less than half a century after the invention of the calculus. For one of Euler's type of genius no time could have been better. Rich in his instinct for the mathematical character of a physical phenomenon, he was seemingly inexhaustible in his power for formulation and invention. With magnificent skill as a mathematical technician he combined an enduring memory for details in all their minuteness and a versatility of interests beyond that of any other. He has been called "the myriad-eyed Euler," and the appellation is fit, notwithstanding the fact that he worked during the last seventeen years of his life in total blindness.

It is distinctive of any epoch in which mathematical fields for new discovery are at hand, that the explorer is prone to hasten on, even when the advance must be made over ground that is only dubiously secure. The attention is then more readily given to the finding of new formulas and solutions, than to the slower and more painstaking reflections which are requisite to the construction of rigorous proofs. We say such a time is one of mathematical *formalism*. Euler's was such a time, and Euler was a characteristic formalist. His work, therefore, radiant

of ingenuity and imagination as it is, is also often naive and logically uncompelling. Astonishingly often he breaks out new paths and shows us the way, but fails to prove that the way is there with the conclusiveness that we demand in our time.

I propose, in the following pages, to present one of Euler's derivations which is typical in all these respects, and which shows his genius for solving a difficult problem by elementary means. My presentation will not be a facsimile. I have "edited" it rather drastically, but, I believe, without impairing either its essential spirit or content. The original is to be found in a longish memoir written in Latin [1], where it is imbedded in considerable discursive material. It has been noted that by this derivation Euler came to the brink of one of the greatest of mathematical discoveries, one that was made years later by Fourier. I shall point this out in its proper place.

There are two mathematical relations that become familiar even to college freshmen, for which we are indebted to Euler, and which have been called the "heroes" of his works. These are, namely, the definition of the exponential function which is basic to the elementary calculus

$$e^x = \lim_{n\to\infty}\left\{1+\frac{x}{n}\right\}^n, \tag{1}$$

and the cardinal formula of analytic trigonometry

$$e^{ix} = \cos x + i \sin x. \tag{2}$$

In what follows, these heroes may be seen in action.

The problem Euler set himself is the following one:

Given an arbitrary periodic function $f(x)$ with the period 1, *to find a function $y(x)$ which is changed by the amount $f(x)$ when the variable is increased from x to $x+1$.*

In symbols, this is a call for a solution of the *difference equation*

$$y(x+1) = y(x) + f(x). \tag{3}$$

Euler bases his attack upon the formula

$$F(a+h) = F(a) + F'(a)h + \frac{1}{2!}F''(a)h^2 + \frac{1}{3!}F'''(a)h^3 + \cdots$$

namely, upon the Taylor's series, which had been published in 1715. With F, a and h, taken respectively to be y, x and 1, this gives

$$y(x+1) = y(x) + y'(x) + \frac{1}{2!}y''(x) + \frac{1}{3!}y'''(x) + \cdots,$$

and this, by comparison with the equation (3), shows that the function $y(x)$ sought is one for which

$$y'(x) + \frac{1}{2!}y''(x) + \frac{1}{3!}y'''(x) + \cdots = f(x). \tag{4}$$

In other words, what is wanted is a solution of the equation (4), an ordinary linear differential equation of infinite order.

It is common practice in dealing with ordinary differential equations to symbolize the kth derivative of y by $D^k y$. The equation (4) is thereby given the form

$$\left\{D + \frac{1}{2!}D^2 + \frac{1}{3!}D^3 + \cdots\right\} y = f(x),$$

and this is unambiguously indicated by writing it

$$\{e^D - 1\}y = f(x), \tag{5}$$

since for any value of z we have the power series expansion

$$e^z = 1 + z + \frac{1}{2!}z^2 + \frac{1}{3!}z^3 + \cdots.$$

Euler, therefore, seeks a solution to the symbolically expressed equation (5).

His appeal at this point is to his hero relation (1). In accordance with that the equation (5) is, in other terms,

$$\lim_{n\to\infty} \left\{\left(1 + \frac{D}{n}\right)^n - 1\right\} y = f(x).$$

He is led by this to explore the ground by considering the "approximating" equation

$$\left\{\left(1 + \frac{D}{n}\right)^n - 1\right\} u = f(x), \tag{6}$$

in which n is a specific but unspecified integer. He finds it convenient to take this integer to be odd, and since the purpose is only one of exploration, nothing is to be lost by taking it so. Euler's thought in this connection was, of course, the following. Whatever the integer n, the equation (6) is linear and has constant coefficients. It is therefore of an elementary type whose solution can be explicitly written out. By observing how that solution "matures" when n is permitted to become infinite, one should be able to discern the character of a solution of the equation (5).

The common method of dealing with an equation (6) is, first, to write it as

$$u = \frac{1}{\left(1 + \frac{D}{n}\right)^n - 1} f(x),$$

then, to treat the multiplier of $f(x)$ in this as an algebraic fraction in D, and as such to resolve it into partial fractions:

$$\frac{1}{\left(1+\frac{D}{n}\right)^n - 1} = \sum_{j=1}^{n} \frac{c_j}{D - a_j}. \tag{7}$$

Thereby the equation (6) is made to appear as

$$u = \sum_{j=1}^{n} \frac{c_j}{D - a_j} f(x),$$

and a solution of it is accordingly

$$u(x) = \sum_{j=1}^{n} u_j(x), \tag{8}$$

where the terms $u_j(x)$ are respectively solutions of the equations

$$u_j(x) = \frac{c_j}{D - a_j} f(x),$$

which is to say of the first order differential equations

$$\{D - a_j\} u_j = c_j f(x). \tag{9}$$

To carry this program out the constants a_j and c_j must be determined. We proceed to do that.

The constants a_j are the values of D for which the denominator of the left-hand member of the equation (7) is zero, namely are the roots of the equation

$$\left(1+\frac{D}{n}\right)^n = 1. \tag{10}$$

Now since $\cos 2\pi x$ and $\sin 2\pi x$ have the values 1 and 0 respectively when x has an integral value k, Euler's relation (2) shows that

$$e^{2k\pi i} = 1,$$

and hence that the equation (10) can be written as

$$\left(1+\frac{D}{n}\right)^n = e^{2k\pi i}.$$

On dividing the exponents by n and solving for D, we find as the roots the values

$$n\{e^{2k\pi i/n} - 1\}.$$

Precisely n of these are distinct. We may label them thus:

$$
\begin{aligned}
a_1 &= 0, \\
a_{2j} &= n\left\{e^{2j\pi i/n} - 1\right\}, \qquad j = 1, 2, \cdots, \frac{n-1}{2}. \\
a_{2j+1} &= n\left\{e^{-2j\pi i/n} - 1\right\},
\end{aligned} \tag{11}
$$

The method for determining the constants c_k is a standard one. Because the relation (7) is conceived of as an identity in D, it must continue to hold when D is varied in any way. If we first multiply it by $D-a_k$ and then let D approach a_k as a limit, every term in the sum except that for which $j=k$ approaches zero. Thus we infer that

$$
\lim_{D\to a_k} \frac{D - a_k}{\left(1+\frac{D}{n}\right)^n - 1} = c_k.
$$

Hereby c_k is given as the limiting value of an indeterminate fraction of the 0/0 type. The method to evaluate this is to replace the numerator and the denominator of the fraction by their derivatives with respect to D before proceeding to the limit. Thus

$$
c_k = \lim_{D\to a_k} \frac{1}{\left(1+\frac{D}{n}\right)^{n-1}} = \frac{1}{\left(1+\frac{a_k}{n}\right)^{n-1}}.
$$

Since a_k is a root of the equation (10), we can write the result

$$
c_k = 1 + \frac{a_k}{n}. \tag{12}
$$

With this, Euler's exploratory base is completely laid. The values (11) and (12) make the equations (9) specific, and in terms of their solutions the function $u(x)$ fulfilling the equation (6) is given by the formula (8). He is ready now to let n become infinite, and to observe the resultant metamorphosis of the solution into one for the equation (5).

For each k

$$
\lim_{n\to\infty} e^{2k\pi i/n} = e^0 = 1.
$$

The formulas (11) for a_{2j} and a_{2j+1} are therefore indeterminate of the $\infty \cdot 0$ type as $n\to\infty$. On setting $n=1/s$ we may evaluate their limits thus

$$
\begin{aligned}
\lim_{n\to\infty} a_{2j} &= \lim_{s\to 0} \frac{e^{2j\pi i s} - 1}{s} = 2j\pi i, \\
\lim_{n\to\infty} a_{2j+1} &= \lim_{s\to 0} \frac{e^{-2j\pi i s} - 1}{s} = -2j\pi i.
\end{aligned} \tag{13}
$$

From (12), therefore, $\lim c_k = 1$, and thus the formulas (8) and (9) transform into

$$y(x) = \sum_{j=1}^{\infty} y_j(x), \tag{14}$$

with

$$\begin{aligned} \{D - 2j\pi i\} y_{2j} &= f(x), \\ \{D + 2j\pi i\} y_{2j+1} &= f(x). \end{aligned} \tag{15}$$

We can rewrite (14) as

$$y(x) = y_1(x) + \sum_{j=1}^{\infty} \{y_{2j}(x) + y_{2j+1}(x)\}, \tag{16}$$

and integrate the equations (15) to obtain

$$y_{2j}(x) = \int_a^x f(t) e^{2j\pi i(x-t)} dt,$$

$$y_{2j+1}(x) = \int_a^x f(t) e^{-2j\pi i(x-t)} dt,$$

where a is any convenient constant.

Then, since Euler's relation (2) shows that

$$e^{2j\pi i(x-t)} + e^{-2j\pi i(x-t)} = 2 \cos 2j\pi(x - t),$$

we find that the formula (16) shapes up to

$$y(x) = \int_a^x f(t) dt + \sum_{j=1}^{\infty} 2 \int_a^x f(t) \cos 2j\pi(x - t) dt. \tag{17}$$

This is the elegant formula that Euler has derived for a solution of the difference equation (3).

Imaginative the deduction certainly is. As a proof of its conclusion we find it deficient, especially in a point of reasoning which was generally overlooked in Euler's time. The issue lodges in the passage from a finite n to the infinite, and was clearly recognized only about a century later. In the present instance it is involved in the passage from the equations (8) and (9) to the corresponding ones (14) and (15). This passage was based upon the limiting relations (13).

The correctness of the evaluations (13) for any specific integer j is unquestionable. This means that the values a_{2j} and a_{2j+1} can be made to approximate with greater and greater accuracy $2j\pi i$ and $-2j\pi i$ respectively, and hence that $u_j(x)$ approximates $y_j(x)$ more and more closely, as n is taken larger and larger. But by taking n so, more and more terms are brought into the sum (8). To make these terms good approximations in their turn, more terms again must be

brought in, and the process is thus unending. It is therefore not possible to conclude that because any term $u_j(x)$ can be made to approach $y_j(x)$ as a limit, therefore the whole sum of the $u_j(x)$ approaches the infinite series of the $y_j(x)$.

The oversight of this important logical point mars many mathematical works of the eighteenth century. This does not signify that the works are valueless or uninteresting. In many cases the reasoning can be made rigorous, and even aside from that the deductions often show clearly the character of important presumptive results. Inconclusive, or even erroneous, reasoning by no means implies a wrong conclusion. A quotation on this point is appropos [2]. In his *The Development of Mathematics*, E. T. Bell says:

"This raises an extremely interesting question: how did the master analysts of the eighteenth century—the Bernoullis, Euler, Lagrange, Laplace—contrive to get consistently right results in by far the greater part of their work in both pure and applied mathematics . . . [when what they] mistook for valid reasoning . . . is now universally regarded as unsound? No short answer is possible; but history shows that frequently the essential usable part of a mathematical doctrine is grasped intuitively long before any rational basis is provided for the doctrine itself. The creative mathematicians between Newton and Cauchy obtained mostly correct results . . . because . . . they had instinctively apprehended the self-consistent part of their mathematics."

He continues then, lest we rate our own sophistication too highly:

"Just as no short answer can dispose of our predecessor's good fortune, so none can dispose of ours. Like them, we consistently get meaningful results, although we realize that there is much obscurity in the foundations of our own analysis."

Criticisms of Euler's deduction, other than the one we have enlarged upon, have by no means been wanting. Such were made even contemporaneously by d'Alembert. Nevertheless the formula (17) is valid when the function $f(x)$ is possessed of certain properties. The ways by which that has been most generally established have taken the formula itself as their point of departure, ignoring how it was come by, and shown that it does give a solution of Euler's problem.

It has been observed [3]—and we include this observation by way of concluding this presentation—that the formula (17) lies only a step from one of the greatest deductions of mathematical analysis. Euler did not take this step, most probably because he had no motivation for doing so. It remained for Fourier to do so more than half a century after the date of Euler's work, namely in the early nineteenth century.

If in the formula (17) we replace x by $x+1$, and observe that the factor $\cos 2j\pi(x-t)$ is not thereby changed, we see that

$$y(x+1) = \int_a^{x+1} f(t)dt + \sum_{j=1}^{\infty} 2\int_a^{x+1} f(t)\cos 2j\pi(x-t)dt,$$

and hence that

$$(18) \qquad y(x+1) - y(x) = \int_x^{x+1} f(t)dt + \sum_{j=1}^{\infty} 2\int_x^{x+1} f(t)\cos 2j\pi(x-t)dt.$$

By (3) the left-hand member of this equation is $f(x)$. In the right-hand member each integral has an integrand that is periodic with the period 1, and an interval of integration that is precisely such a period in length. The interval of integration can therefore be replaced by any other one of the same length, in particular by that from 0 to 1. Thus one finds that

$$f(x) = \int_0^1 f(t)dt + \sum_{j=1}^{\infty} 2\int_0^1 f(t)\cos 2j\pi(x-t)dt.$$

When the cosine factors in this are expanded by the trigonometric addition formula, the result reads

$$(19) \qquad f(x) = \frac{A_0}{2} + \sum_{j=1}^{\infty} \{A_j \cos 2j\pi x + B_j \sin 2j\pi x\},$$

with

$$(20) \qquad \begin{aligned} A_j &= 2\int_0^1 f(t)\cos 2j\pi t dt, \\ B_j &= 2\int_0^1 f(t)\sin 2j\pi t dt. \end{aligned}$$

This is the Fourier expansion of the arbitrary function $f(x)$.

References

1. Euleri, L., De serierum determinatione seu nova methodus inveniendi terminos generalis serierum, Opera Omnia, Series Prima, xiv, p. 510.

2. Bell, E. T., The Development of Mathematics, New York, 1940.

3. Burkhardt, H., Entwickelungen nach oscillirenden Functionen und Integration der Differentialgleichungen der mathematischen Physik. Jahresbericht der Deutschen Mathematiker-Vereinigung, 10, 2, 1, Leipzig, 1908, p. 50.

QUANTUM MECHANICS AND HILBERT SPACE

GEORGE W. MACKEY, Harvard University

1. Introduction. This expository article is written in the spirit of the various "What is a ———?" articles that have appeared in this MONTHLY from time to time. It is an attempt to give mathematicians who have not studied quantum mechanics some idea of what it is about and to give those who have studied it only from the point of view of the physicists some idea of how it may be formulated† in a mathematically precise fashion. We shall avoid the difficulties that still plague the advanced portions of the subject by dealing only with non-relativistic quantum mechanics.

2. The nature of quantum mechanics. The basic difference between classical mechanics and quantum mechanics is that the $2n$-tuple of real numbers which characterizes the state of a system in classical mechanics is replaced in quantum mechanics by a family of probability measures on the real line. The $2n$ first-order differential equations in $2n$-space which describe how the state of the classical system changes in time are replaced by a differential equation in an infinite-dimensional space which describes how these probability measures change with time. Under certain circumstances the probability measures are so highly concentrated (*i.e.*, assign a probability so close to one to such a short interval) that they behave like numbers. The resulting system of numbers then changes approximately according to the laws of classical mechanics. In this sense classical mechanics is a limiting case of quantum mechanics.

The passage from classical mechanics to the more exact quantum mechanics becomes necessary when one can no longer ignore the way in which different measurements interfere with one another—for example, when dealing with the submicroscopic world. Even in dealing with familiar large-scale phenomena this passage would become necessary if one were to make measurements to a much higher degree of accuracy than is now possible.

The reason that the above-mentioned interference forces a retreat from numbers to probability measures may be explained as follows. The act of measuring the value of an observable A may change the system in such a way that a measurement of B immediately following the measurement of A will not agree with a measurement of B immediately preceding the measurement of A. Moreover the change in the result of measuring B will be of an unpredictable nature. As analysis of the situation shows, this circumstance precludes the possibility of assigning any meaning to the statement "A and B have the values a and b at the time t." On the other hand it does not preclude the possibility of assigning

† While the formulation given here is based essentially on that to be found in von Neumann's classic book, Mathematische Grundlagen der Quantenmechanik, it differs from it in a number of respects. It is the writer's personal way of looking at the foundations of quantum mechanics and cannot be claimed to be the standard mathematical view if indeed there is such. For another approach the reader is referred to Chapter I of some recent University of Chicago notes by Irving Segal.

a meaning to the statement "At the time t measurements of A and B are statistically distributed with probability measures† α and β. This is because a statistical distribution is determined by making measurements on a sample containing a large number of replicas of the system under consideration and one may use distinct samples in determining α and β.

3. General mathematical formulation. Let $\mathfrak{O}$ denote the set of all "observables" associated with a given physical system. By a *state* of our system we shall mean a *possible* simultaneous set of statistical distributions for the "observables" in $\mathfrak{O}$. We denote the set of all states by $\mathfrak{S}$. It will not be necessary for us to analyze the notions of state and observable. From a purely mathematical point of view $\mathfrak{O}$ is an abstract set and each member of $\mathfrak{S}$ is a function defined on $\mathfrak{O}$ and having values in the set $\mathfrak{M}$ of all probability measures on the line. The system $\mathfrak{O}$, $\mathfrak{S}$ determines completely what we may call the statics of the system. To determine the dynamics we must specify how members of $\mathfrak{S}$ change with time; that is we must determine for each real t the one-to-one transformation L_t of $\mathfrak{S}$ onto $\mathfrak{S}$ such that $L_t(\phi)$ is the state of the system a time t after it was in the state ϕ. Of course the family L_t need not be given directly in integrated form, but may be determined only by giving some sort of differential equation in t.

It should be noted that our formulation is still broad enough to include classical mechanics. In that case $\mathfrak{O}$ is the set of all sufficiently regular real-valued functions defined on classical phase space $\mathfrak{P}$ and the general member of $\mathfrak{S}$ is defined by a probability measure on $\mathfrak{P}$. The function from $\mathfrak{O}$ to $\mathfrak{M}$ defined by the probability measure α in $\mathfrak{P}$ takes the function f on $\mathfrak{P}$ into the probability measure $E \to \alpha(f^{-1}(E))$. The L_t are obtained by integrating the classical equations of motion on $\mathfrak{P}$ and using the resulting one parameter group of one-to-one transformations of $\mathfrak{P}$ onto itself to induce a corresponding group of one-to-one transformations of $\mathfrak{S}$ onto itself. Actually, of course, classical mechanics is usually concerned only with the special states defined by those probability measures which are concentrated in a single point of $\mathfrak{P}$. Indeed the motion of these determines the motion of all of the others and it is for this reason that classical mechanics‡ need not concern itself with measures as such.

Given two states of a physical system one can consider a third state in which the system is in the first state with probability p and in the second with probability $1-p$. This leads us to make the following assumption about $\mathfrak{S}$. If ϕ_1 and ϕ_2 are in $\mathfrak{S}$ and $0 \leqq p \leqq 1$ then $p\phi_1+(1-p)\phi_2$ is also in $\mathfrak{S}$. A state ϕ which can be so constructed from two other states is said to be a *mixed state*. A state which cannot be so constructed is said to be a *pure state*. When $\mathfrak{O}$ and $\mathfrak{S}$ are as in the preceding paragraph the pure states are obviously just those defined by probability measures concentrated in single points of $\mathfrak{P}$. Physically speaking, a pure

† See the foundations of probability, by P. R. Halmos, this MONTHLY vol. 51, 1944, pp. 493-510.

‡ The word "need" here must be interpreted in a suitably restricted sense. Of course, it may still be quite important to discuss the motion of measures which are not concentrated in points—for example, in classical statistical mechanics.

state is one in which one cannot concentrate the probability measure for any observable without spreading out the probability measure for another. The characteristic features of quantum mechanics arise from the fact that the pure states are not trivially pure as they are in classical mechanics—for each pure state there are many observables with highly unconcentrated probability measures.

If f is a real-valued function of a real variable and A is an observable, there is an obvious physical sense in which one can construct an observable $f(A)$. A measurement of $f(A)$ is simply a measurement of A, giving the value a say, followed by a computation of $f(a)$. This leads us to make the following further assumption about the mathematical system $\mathfrak{O}$, $\mathfrak{S}$. For each $A\in\mathfrak{O}$ and each real Borel function f on the real line there exists $B\in\mathfrak{O}$ such that for every Borel set E and each $\phi\in\mathfrak{S}$ $\phi(B)(E)=\phi(A)(f^{-1}(E))$. If we assume, as we shall, that no two distinct members of $\mathfrak{O}$ are carried into the same probability measure by all members of $\mathfrak{S}$, then B is uniquely determined by f and A. We denote it by $f(A)$.

4. The transition from observables to questions. Let us call an observable A a *question* if for every state ϕ the probability measure $\phi(A)$ is concentrated in the two points 0 and 1. It is not difficult to prove that A is a question if and only if $A^2=A$. If A is any observable and E is any Borel subset of the real line let ψ_E be the characteristic function of E; that is the function which is one for $x\in E$ and zero for x not in E. Then $\psi_E(A)$ is always a question which we shall denote by Q_E^A. The family of questions Q_E^A obtained by holding A fixed and letting E vary through the Borel subsets of the real line clearly determines the observable A. When Q is a question and ϕ is a state the probability measure $\phi(Q)$ is completely determined by its value at the set consisting of one alone. We shall call this value $m_\phi(Q)$. Since $m_\phi(Q_E^A)=m_\phi(\psi_E(A))=\phi(A)(\psi_E^{-1}(\{1\}))=\phi(A)(E)$ we see that the state ϕ is completely determined by the values of m_ϕ at the questions in $\mathfrak{O}$.

Let us denote the set of all questions in $\mathfrak{O}$ by q. We partially order q by setting $Q_1\leqq Q_2$ whenever $m_\phi(Q_1)\leqq m_\phi(Q_2)$ for all states ϕ. If Q_1 and Q_2 are questions such that $Q_1\leqq 1-Q_2$, we say that Q_1 and Q_2 are mutually exclusive and write $Q_1\perp Q_2$. If $Q_1, Q_2, \cdots$ are questions such that $Q_i\perp Q_j$ for $i\neq j$ there clearly exists at most one question Q such that $m_\phi(Q)=m_\phi(Q_1)+m_\phi(Q_2)+\cdots$ for all $\phi\in\mathfrak{S}$. We assume that Q always exists and denote it by $Q_1+Q_2+\cdots$. It is not difficult to show† that Q is the least question R such that $R\geqq Q_j$ for all j.

If A is an observable then the mapping $E\to Q_E^A$ clearly has the following properties: (1) $E\cap F=0$ implies $Q_E^A\perp Q_F^A$, (2) $Q_{\{0\}}^A=0$ and $Q_{[-\infty,\infty]}^A=1$,‡ (3) $E=E_1\cup E_2\cup\cdots$ where $E_i\cap E_j=0$ for $i\neq j$ implies that $Q_E^A=Q_{E_1}^A+Q_{E_2}^A+\cdots$. If $E\to Q_E$ is any mapping of Borel sets in the line into questions which has the three properties just listed we shall call it a *question-valued measure*. Let $\bar{\mathfrak{O}}$ denote the set of all question-valued measures. For each $\phi\in\mathfrak{S}$ and $Q\in\bar{\mathfrak{O}}$ let

† We are indebted to R. V. Kadison for this remark. In our original manuscript the fact in question appears as an added assumption.

‡ 1 is the unit constant observable, that is the observable A such that $\phi(A)$ is concentrated in the point 1 for all states ϕ.

$\phi(Q)(E)=m_\phi(Q_E)$. If we identify the members A of $\mathfrak{O}$ with their associated question-valued measures Q^A so that we may consider $\mathfrak{O}$ as a subset of $\bar{\mathfrak{O}}$, then the preceding definition amounts to an extension of each ϕ from $\mathfrak{O}$ to $\bar{\mathfrak{O}}$. The system $\bar{\mathfrak{O}}$, $\mathfrak{S}$ then satisfies all axioms already laid down for $\mathfrak{O}$, $\mathfrak{S}$, and in addition the following: Every question-valued measure is the question-valued measure associated with some observable. We shall assume that we have added the observables in $\bar{\mathfrak{O}}-\mathfrak{O}$ to our system, that is, that our system $\mathfrak{O}$, $\mathfrak{S}$ satisfies the axiom just enunciated. Thus $\mathfrak{O}$ and $\mathfrak{S}$ can be reconstructed as soon as we are given q and the m_ϕ. It is clear that we could have taken q and the m_ϕ as our starting-point instead of $\mathfrak{O}$ and $\mathfrak{S}$.

The m_ϕ clearly have the following properties: (1) If $Q_1, Q_2, \cdots$ are questions such that $Q_i \perp Q_j$ for $i \neq j$ then $m_\phi(Q_1+Q_2+\cdots)=m_\phi(Q_1)+m_\phi(Q_2)+\cdots$, (2) $m_\phi(1)=1$. We shall call any function from the questions to the positive real numbers which has these two properties a *measure on the questions*. It is tempting to add the assumption that every measure on the questions is the m_ϕ for some state ϕ. Like the corresponding assumption for question-valued measures it can be attained by enlarging the system. However, the physical significance of states and observables differ in such a way that one is much more willing to make the first enlargement than the second. There is a sense in which one can construct observables from questions, but we have to be content with the states provided by nature. If we did add this assumption (and for all one knows to the contrary it holds for the system $\mathfrak{O}$, $\mathfrak{S}$ actually in use for quantum mechanics) it would follow that the pair $\mathfrak{O}$, $\mathfrak{S}$ is completely determined by the partially ordered set q of all questions and the "complementation operation" $Q \to 1-Q$. In any event it is natural to call this complemented partially ordered set the *logic*† of our system. The statics of the system is determined by its logic and a certain convex set of measures on the questions.

5. The logic of quantum mechanics. We have not yet added enough axioms to exclude classical mechanics. If we interpret "sufficiently regular" in our description of the $\mathfrak{O}$, $\mathfrak{S}$ for classical mechanics to mean being a Borel function, then the questions are in a natural one-to-one correspondence with the Borel subsets of phase space. The logic of classical mechanics is a Boolean algebra—the Boolean algebra of all Borel subsets of phase space. We see in particular that in the case of classical mechanics our questionable axiom is indeed satisfied. Every measure on the questions corresponds to a state. It is interesting to note that any two classical mechanical systems have isomorphic logics. This follows from a known set-theoretical result according to which there exists a one-to-one Borel set-preserving map between any two separable complete metric spaces.

We single out the system of quantum mechanics not by adding more general axioms but by stating explicitly what its logic is and which measures on the

† *Cf.* The logic of quantum mechanics, by G. Birkhoff and J. von Neumann, Ann. Math., vol. 37, 1936, pp. 823–843.

questions are associated with states. As in the case of classical mechanics, all quantum mechanical systems have (to within isomorphism) the same logic. This logic is the partially ordered set of all closed subspaces of a separable infinite dimensional Hilbert space (see the next section for a definition of Hilbert space), complementation being passage from a subspace to its orthogonal complement.

It would, of course, be very interesting if this structure for q and $Q \to 1-Q$ could be deduced from a set of physically meaningful and plausible axioms. Such a deduction is not at present available. It has simply been found that physical systems behave as if q had the indicated structure. For us it is an axiom. To mathematicians familiar with lattice theory this axiom is not so artificial and arbitrary as it may seem at first sight to others. The lattice of all subspaces of a vector space is one of the principal examples of a partially ordered set which is not a Boolean algebra but nevertheless has many regularity properties. Moreover, one can show that if the lattice of all closed subspaces of a Banach space has an order inverting involution with the properties of $Q \to 1-Q$, then this Banach space has an equivalent norm under which it is a Hilbert space. We shall postpone our description of $\mathcal{S}$ until after the exposition of basic Hilbert space theory to which we propose to devote the next two sections.

6. Definition and elementary properties of Hilbert space. A Hilbert space $\mathfrak{H}$ is a vector space over the complex numbers in which there is given a complex-valued function of two variables (ϕ, ψ) such that: (1) For fixed ψ, (ϕ, ψ) is a linear function of ϕ, (2) $(\phi, \psi) = \overline{(\psi, \phi)}$, (3) $(\phi, \phi) > 0$ unless $\phi = 0$. (4) Under $\rho(\phi, \psi) = \sqrt{(\phi-\psi, \phi-\psi)}$, $\mathfrak{H}$ is a complete metric space. The complex number (ϕ, ψ) is called the inner product of ϕ and ψ, and $\sqrt{(\phi, \phi)}$ is called the norm of ϕ and written $\|\phi\|$.

The two most familiar examples of Hilbert spaces are the following: (I) $\mathfrak{H}$ is the set of all sequences $c_1, c_2, \cdots$ of complex numbers such that $|c_1|^2 + |c_2|^2 + \cdots < \infty$, and the inner product of $c_1, c_2, \cdots$ with $d_1, d_2, \cdots$ is $c_1\bar{d}_1 + c_2\bar{d}_2 + \cdots$, (II) $\mathfrak{H}$ is the set of all complex-valued Lebesgue measurable square-integrable functions on the interval $[a, b]$ $(-\infty \leqq a < b \leqq \infty)$, functions equal almost everywhere being identified, with $(f, g) = \int_a^b f(x)\bar{g}(x)dx$. On modifying (II) by passing from $[a, b]$ to a general measure space one gets a class of examples including both (I) and (II). It is important to note that except for differences in "dimension" all Hilbert spaces are abstractly equivalent—the dimension of a Hilbert space being defined as the maximum number of elements in a set $\{\phi_\alpha\}$ such that $(\phi_\alpha, \phi_\beta) = 0$ for $\alpha \neq \beta$ and $(\phi_\alpha, \phi_\alpha) = 1$. Actually any two such sets which are maximal in the sense that they cannot be enlarged and which lie in the same Hilbert space have the same cardinal number. Such a maximal set is said to be an orthonormal basis for the space. If $\mathfrak{H}$ has dimension $\aleph_0$, as do examples (I) and (II) above, and $\phi_1, \phi_2, \cdots$ is an orthonormal basis, then every ϕ in $\mathfrak{H}$ is uniquely representable in the form $c_1\phi_1 + c_2\phi_2 + \cdots$, where $|c_1|^2 + |c_2|^2 + \cdots < \infty$. The mapping $\phi \to \{c_1, c_2 \cdots\}$ is then an inner product preserving linear map of $\mathfrak{H}$ on example (I). Practically the same considerations lead to a correspond-

ing proof for higher dimensional spaces. A separable Hilbert space always has dimension $\aleph_0$ or less.

A closed subspace of a Hilbert space is by definition a linear subspace which is closed in the metric. Every closed subspace is a Hilbert space in its own right. If M is a closed subspace then $M^\perp$, the orthogonal complement of M, is by definition the set of all ψ such that $(\phi, \psi)=0$ for all $\phi \in M$. It can be shown that $M^{\perp\perp}=M$ for all M, and that every $\phi \in \mathfrak{C}$ is uniquely of the form $\phi_1+\phi_2$, where $\phi_1 \in M$ and $\phi_2 \in M^\perp$.

7. Linear operators in Hilbert space. Let $\mathfrak{C}$ be a Hilbert space. Let T be a linear function mapping $\mathfrak{C}$ into $\mathfrak{C}$. T is continuous if and only if it is bounded, in the sense that $\|T(\phi)\|/\|\phi\|$ is bounded for $\phi \neq 0$. The least upper bound of these numbers $\|T(\phi)\|/\|\phi\|$ is called the norm $\|T\|$ of T. For each bounded linear operator T there is a unique bounded linear operator T^* such that $(T(\phi), \psi) = (\phi, T^*(\psi))$ for all ϕ and ψ in $\mathfrak{C}$. T^* is called the adjoint of T. If $T=T^*$, T is said to be *self-adjoint*. If $T^*T=TT^*=I$ (where I is the identity operator) then T is said to be *unitary*. T is unitary if and only if it is one-to-one, has all of $\mathfrak{C}$ for its range and preserves the inner product. The unitary operators are thus the automorphisms of $\mathfrak{C}$. Let M be a closed subspace of $\mathfrak{C}$. For each $\phi \in \mathfrak{C}$ let $P_M(\phi)$ denote the unique $\psi \in \mathfrak{C}$ such that $\phi-\psi$ is in $M^\perp$; that is, let P_M be the "projection" of ϕ on M. Then P_M is a bounded self-adjoint linear operator and $P_M^2=P_M$. Conversely if T is any self-adjoint bounded linear operator such that $T^2=T$, it is easy to show that $T=P_M$ where M is the set of all ϕ such that $T(\phi)=\phi$. For this reason the bounded self-adjoint linear operators T such that $T^2=T$ are called *projections*. The mapping $M \to P_M$ sets up a one-to-one correspondence between the closed subspaces and the projections such that for all M, $P_{M^\perp}=1-P_M$. Let T be any bounded self-adjoint linear operator. Then $(T(\phi), \psi)$ as a function of ϕ and ψ has properties (1) and (2) listed in the definition of an inner product. Moreover it also has the property: $|(T(\phi), \psi)| \leqq \|T\| \cdot \|\phi\| \, \|\psi\|$. We shall call any function $[\phi, \psi]$ which has properties (1) and (2), and is such that $|[\phi, \psi]| \leqq K\|\phi\| \, \|\psi\|$ for some positive real number K, a bounded Hermitian bilinear form. It is a theorem that every bounded Hermitian bilinear form is of the form $(T(\phi), \psi)$ for a uniquely determined self-adjoint bounded linear operator T.

8. Quantum statics. Let $\mathfrak{C}$ be the Hilbert space whose lattice of closed linear subspaces is the logic of our quantum-mechanical system $\mathfrak{S}$, $\mathfrak{s}$ so that we may identify questions in $\mathfrak{S}$ with closed subspaces of $\mathfrak{C}$. Let ϕ be any element of $\mathfrak{C}$ such that $\|\phi\|=1$. Then $M \to (P_M(\phi), \phi)$ is readily verified to be a measure on the questions. More generally if $\phi_1, \phi_2, \cdots$ is any sequence of members of $\mathfrak{C}$ all having norm one and $\gamma_1, \gamma_2, \cdots$ is any sequence of positive real numbers such that $\gamma_1+\gamma_2+\cdots=1$ then $\gamma_1(P_M(\phi_1) \cdot \phi_1)+\gamma_2(P_M(\phi_2), \phi_2)+\cdots$ converges for all M and as a function of M is a measure on the questions. We complete our axioms concerning quantum statics by assuming that the measures on the ques-

tions which define states are just those of the form† $M \to \gamma_1(P_M(\phi_1), \phi_1) + \gamma_2(P_M(\phi_2), \phi_2) + \cdots$. It is easy to show that a state is pure if and only if it is defined by a single element of $\mathfrak{K}$; that is if and only if the associated measure on the questions is of the form $M \to (P_M(\phi), \phi)$. It is also easy to show that ϕ_1 and ϕ_2 define the same pure state if and only if $\phi_1 = e^{i\theta}\phi_2$ for some real θ.

If A is any observable and $E \to Q_E^A$ is the corresponding question-valued measure we may construct what we shall call a "projection-valued measure" by replacing each closed subspace Q_E^A by the projection upon it. We leave the formulation of the definition of projection-valued measure to the reader. We will denote the projection-valued measure thus associated with the observable A by P^A and the projection which it assigns to the Borel set E by P_E^A. Obviously the mapping $A \to P^A$ sets up a one-to-one correspondence between all observables on the one hand and all projection-valued measures on the other. The probability distribution of the observable A in the pure state defined by the Hilbert-space element ϕ is of course just $E \to (P_E^A(\phi), \phi)$—a formula which sums up the whole of quantum statics. In the sequel we shall speak only of pure states and we shall find it convenient to identify them with the Hilbert-space elements which define them.

9. Observables and operators. Let A be an observable which is *bounded* in the sense that $P_E = I$ for some bounded subset E of the real line. Then for all ϕ and ψ in $\mathfrak{K}$, $E \to (P_E^A(\phi), \psi)$ is a countably additive complex-valued set function $\alpha_{\phi,\psi}^A$ with respect to which the Stieltjes integral $\int_{-\infty}^{\infty} x d\alpha_{\phi,\psi}^A(x)$ exists. The resulting function of ϕ and ψ is easily seen to be a bounded Hermitian bilinear form. Hence there exists a bounded self-adjoint linear operator T^A uniquely determined by A such that for all ϕ and ψ in $\mathfrak{K}$, $(T^A(\phi), \psi) = \int_{-\infty}^{\infty} x d\alpha_{\phi,\psi}^A(x)$. Thus every bounded observable A is associated with a unique bounded self-adjoint linear operator. It follows from the celebrated spectral theorem that the converse is true. Every bounded self-adjoint linear operator is derivable in the manner just described from a uniquely determined bounded projection-valued measure, and hence is of the form T^A for a uniquely determined bounded observable A. $A \to T^A$ thus sets up a one-to-one correspondence between the set of all bounded observables in $\mathfrak{O}$ and the set of all bounded self-adjoint linear operators. This correspondence may be extended to one between all observables in $\mathfrak{O}$ and all self-adjoint linear operators by suitably defining the notion of unbounded self-adjoint linear operator. Let T be linear but let its domain be (perhaps) only a dense linear subspace of $\mathfrak{K}$, and let us make no assumptions about its boundedness or continuity. The adjoint of T is then defined as follows: ψ is in the domain

† As indicated earlier, no other measures on the questions are known at present. Indeed, since these words were written, A. M. Gleason has made considerable progress toward showing that there are none. Using his results one can replace our assumption about the specific form of the allowable measures by one which asserts, simply, that these measures are not too badly "non-measurable."

of T^* if and only if $(T(\phi), \psi)$ is continuous as a function of ϕ for all ϕ in the domain of T. $T^*(\psi)$ is then the unique element of $\mathfrak{H}$ such that $(T(\phi), \psi) = (\phi, T^*(\psi))$ for all ϕ in the domain of T. T is said to be self-adjoint whenever $T = T^*$ (equality including having the same domain). When T is bounded and everywhere defined this definition coincides with the one already given. The construction of an unbounded self-adjoint operator from an unbounded projection-valued measure is analogous to the corresponding construction in the bounded case but is complicated by the fact that the relevant bilinear form is not everywhere defined. We shall not give details here. We exploit the one-to-one correspondence between observables and self-adjoint linear operators which we have just described by identifying every observable with its corresponding operator.

Let A be a self-adjoint linear operator in $\mathfrak{H}$ and let P^A be the corresponding projection valued measure. Let G denote the union of all open sets E for which $P^A_E = 0$. Then G itself is an open set and $P^A_G = 0$. The closed complement of the open set G is called the *spectrum* S_A of A. The significance of S_A for A regarded as an observable is obviously the following. In every state the probability measure of A is concentrated in S_A and for every open interval F on the real line which intersects S_A there exists a state in which the probability measure of A is concentrated in F. In this sense S_A is the set of possible values for the observable A. Every closed subset of the real line is the spectrum of some self-adjoint linear operator and hence the set of possible values of some observable.

The set of all $\lambda \in S_A$ such that $P^A_{\{\lambda\}} \neq 0$, where $\{\lambda\}$ denotes the set whose only element is λ, is called the *point spectrum* of A. If ϕ is an element of norm one in the range of $P^A_{\{\lambda\}}$ then it is a state in which the observable A has the value λ with probability one. Conversely if ϕ is a pure state in which A has the value λ with probability one then ϕ is in the range of $P^A_{\{\lambda\}}$. The range of $P^A_{\{\lambda\}}$ coincides with the set of all ϕ for which $A(\phi) = \lambda\phi$; that is with the set of all *proper vectors* of A belonging to the *proper value* λ. The point spectrum of A is always at most denumerable. A is said to have a *pure point spectrum* if $P^A_D = I$ where D is the point spectrum; that is if in every state the probability measure of A is concentrated in D. A has a pure point spectrum if and only if there exists a basis of proper vectors. One of the more spectacular ways in which quantum mechanics differs from classical mechanics is in the occurrence in quantum mechanics of observables with a sizable point spectrum whose classical analogues are continuous functions on phase space.

The identification of observables with self-adjoint operators together with the operation $a \to f(A)$ already defined for observables leads at once to a definition of $f(A)$ whenever f is a Borel function and A is a self-adjoint operator. This definition can of course be formulated without reference to quantum mechanics. When f is a polynomial it agrees with the usual algebraic definition of a polynomial function of an operator. If A_1 and A_2 are bounded self-adjoint linear operators it can be shown that $A_1A_2 = A_2A_1$ if and only if $P^{A_1}_E P^{A_2}_F = P^{A_2}_F P^{A_1}_E$ for all E and F. When A_1 and A_2 are unbounded and consequently not everywhere defined this condition on the projection-valued measures is taken as the defini-

tion of the commutativity of the operators. With this definition it can be shown that A_1 and A_2 commute if and only if there exists a third self-adjoint operator A_3 and Borel functions f_1 and f_2 such that $A_1=f_1(A_3)$ and $A_2=f_2(A_3)$. There is thus an obvious sense in which observables which commute when considered as operators are "simultaneously measurable". The circumstance which prevents our assigning an obvious physical sense to $f(A_1, A_2)$, where f is a Borel function of two variables and A_1 and A_2 are observables, is the possible lack of "simultaneous measurability" of A_1 and A_2. It can be shown that whenever A_1 and A_2 are self-adjoint linear operators for which there is a mapping $f\rightarrow f(A_1, A_2)$ having reasonable properties, then A_1 and A_2 commute. In this sense commutativity of A_1 and A_2 considered as operators is equivalent to simultaneous observability of A_1 and A_2 considered as observables.

10. Quantum dynamics. We have yet to study the nature of the one-parameter group of transformations $t\rightarrow L_t$ of $\mathfrak{S}$ onto $\mathfrak{S}$ which describes the way in which our system $\mathfrak{O}$, $\mathfrak{S}$ changes with time. Our basic dynamical assumption is that each L_t is defined by a unitary transformation U_t of $\mathfrak{H}$ onto $\mathfrak{H}$, and that L_t depends upon t in such a manner that $|(U_t(\phi), \psi)|$ is continuous in t for all ϕ and ψ in $\mathfrak{H}$. Each U_t is determined by L_t up to multiplication by a complex number of modulus one. It turns out that one can choose the arbitrary constants so that $U_{t_1+t_2}=U_{t_1}U_{t_2}$ for all t_1 and t_2, and so that $(U_t(\phi), \psi)$ is itself continuous. Moreover, the fact that L_t is of the given form may be deduced from quite general hypotheses about the action of the L_t on $\mathfrak{S}$. The continuous one-parameter group $t\rightarrow U_t$ is, of course, not quite uniquely determined by the L_t. If c is any real number then $t\rightarrow U_t$ and $t\rightarrow e^{-itc}U_t$ come from the same L_t's and define the same dynamics. This however is the extent of the ambiguity.

Let H be any self-adjoint linear operator. We define e^{-itH} for all real t as $\cos(tH)-i\sin(tH)$. Since $\sin x$ and $\cos x$ are bounded functions, e^{-itH} is a bounded linear operator even when H is not bounded. As a matter of fact, it is unitary, and $t\rightarrow e^{-itH}$ is a continuous one-parameter unitary group of operators. According to a fundamental theorem of M. H. Stone, every continuous one-parameter group of unitary operators is of this form, with a uniquely determined self-adjoint H. Thus $H\rightarrow U_t$, where $U_t=e^{-itH}$, sets up a one-to-one correspondence between the self-adjoint operators and the continuous one-parameter unitary groups. If $t\rightarrow U_t$ determines the dynamics of our system and $U_t=e^{-itH}$, then the self-adjoint operator H also determines the dynamics of our system. Since $e^{-itc}e^{-itH}=e^{-it(c+H)}$, the ambiguity in U produces a certain ambiguity in H as well. H is determined only up to an additive constant operator by the L_t. We shall ignore this ambiguity and suppose a particular H chosen once and for all. We shall call H the *dynamical operator* of the system.

If ϕ is in the domain of H and is of norm one then the variable pure state $e^{-itH}(\phi)=\phi_t$ satisfies the differential equation: $d\phi_t/dt=-iH(\phi_t)$. This is the abstract form of what is called Schrödinger's equation, and plays the role in quantum mechanics played by Hamilton's equations in classical mechanics. It

is a first-order differential equation whose solutions are the trajectories of the pure states.

Let ϕ be a proper vector of H having norm one and corresponding proper value λ. Then $e^{-itH}(\phi) = e^{-it\lambda}\phi$, and thus is the same state for all time t. In other words, the state defined by ϕ is constant in time. It is what is called a *stationary state*. Conversely every stationary state ϕ is a proper vector of H. The existence of pure states which are stationary is a feature of quantum mechanics which has no counterpart in classical mechanics. Of course, a given H need not have any proper vectors, but those which arise most commonly in practice have a good many.

In classical mechanics certain functions on phase space are constant on the trajectories and are called integrals. It is natural to call a quantum mechanical observable an *integral* if its probability-measure in every state ϕ is constant in time. An easy calculation shows that the observable A is an integral if and only if it commutes with H when regarded as an operator. Now H itself is a self-adjoint operator and as such is an observable; moreover this observable is an integral. It clearly plays a central role in the theory. As we shall see, it is a constant multiple of an observable which is the quantum-mechanical analogue of the energy integral of classical mechanics. Thus the stationary states of a quantum-mechanical system are just those in which the energy has a definite value with probability one and each stationary state is associated with a definite value of the energy.

11. The quantum mechanics of n "*interacting particles*". The discussion so far has been rather general and abstract. Except in the case of the energy we have given no indications whatever concerning the physical meaning of the observable identified with a particular operator. Moreover, we have not explained how one constructs the quantum-mechanical refinement of a single classical system. We now remedy these deficiencies by exhibiting and discussing a quantum-mechanical system which in the limiting case of highly concentrated probabilities behaves like a system of n classical mass particles moving under mutual central forces.

Consider the Hilbert space of all square-summable functions on Euclidean $3n$ space R^{3n}. Let D denote the formal differential operator

$$f \to \sum_{j=1}^{n} -\frac{1}{\mu_j}\left(\frac{\partial^2 f}{\partial x_j^2} + \frac{\partial^2 f}{\partial y_j^2} + \frac{\partial^2 f}{\partial z_j^2}\right) + Wf,$$

where W is a real function on R^{3n} and the μ_j are positive real numbers. For a wide class of choices of W there is a unique self-adjoint operator H_D whose restriction to the twice differentiable functions coincides with D. We consider the quantum-mechanical system whose Hilbert space is $\mathfrak{L}^2(R^{3n})$ and whose dynamical operator is H_D. For each real-valued Borel function g on R^{3n} let A_g be the self-adjoint operator $f \to fg$—the domain when g is unbounded being defined in the obvious way. The projection-valued measure associated with A_g can be shown to

take the set E into the projection A_ψ where ψ is the characteristic function of $g^{-1}(E)$. Thus the probability measure for A_g in the state f is

$$E \to \int_{g^{-1}(E)} \cdots \int |f|^2 dx_1 \cdots dz_n.$$

Let f be a state such that the mean value $\overline{A}_{x_j} = \int \cdots \int x_j |f|^2 dx_1 \cdots dz_n$ of A_{x_j} exists. A simple calculation shows that

$$\frac{d^2}{dt^2}(\overline{A}_{x_j}) = -\frac{2}{\mu_j}\int \cdots \int |f|^2 \left(\frac{\partial W}{\partial x_j}\right) dx_1 \cdots dz_n = -\frac{2}{\mu_j}\overline{A}_{\frac{\partial W}{\partial x_j}}$$

whenever the relevant derivatives exist. There are of course similar results for the rates of change of the mean values of the A_{y_j} and the A_{z_j}. When f is zero outside of a small neighborhood of a point in R^{3n} the mean values in these formulas are approximately the coordinates of this point and (assuming W to have continuous partial derivatives) the values of the $\partial W/\partial x_j$, $\partial W/\partial y_j$, $\partial W/\partial z_j$ at this point. In other words, when f varies in such a manner that the probability measures of the A_{x_j}, A_{y_j} and A_{z_j} in that state remain highly concentrated, so that the set of them may be represented approximately by a point in R^{3n}, then this point moves so as to satisfy the differential equations

$$\frac{\mu_j}{2}\frac{d^2x_j}{dt^2} = -\frac{\partial W}{\partial x_j}, \quad \frac{\mu_j}{2}\frac{d^2y_j}{dt^2} = -\frac{\partial W}{\partial y_j}, \quad \frac{\mu_j}{2}\frac{d^2z_j}{dt^2} = -\frac{\partial W}{\partial z_j}.$$

These are simply the classical equations of motion for a system of n mass-particles of masses $\frac{1}{2}\mu_j$ and potential energy function W. They are also the classical equations of motion for such a system where the masses are $\frac{1}{2}c\mu_j$, the potential energy function is cW, and c is any positive real number. In classical mechanics masses are determined by the equations of motion only up to a multiplicative constant which is then fixed by arbitrarily choosing a unit of mass. If quantum mechanics had been thought of at the outset one would presumably have made use of the natural unit of mass provided by Schrödinger's equation and defined $\frac{1}{2}\mu_j$ as the mass of the j-th particle. As it is, the mass m_j of the j-th particle is some constant multiple of $\frac{1}{2}\mu_j$. This constant is the same for all systems (for any given units of mass, length and time) and is generally denoted by $\hbar$. $2\pi\hbar$ is the fundamental constant introduced into the old quantum theory by Max Planck and called Planck's constant. Planck's constant is usually denoted by h. In terms of the m_j, the classical potential energy $V = W/\hbar$, and $\hbar$, Schrödinger's equation for our system takes the standard form:

$$-\frac{1}{i}\frac{\partial f}{\partial t} = -\frac{\hbar}{2}\sum_{j=1}^{n}\frac{1}{m_j}\left(\frac{\partial^2 f}{\partial x_j^2} + \frac{\partial^2 f}{\partial y_j^2} + \frac{\partial^2 f}{\partial z_j^2}\right) + \frac{Vf}{\hbar}.$$

Henceforth, we shall suppose V to be the potential energy arising from mutual central forces between the classical particles so that, in particular, there are linear and angular momentum integrals for the limiting classical system.

In $g \rightarrow A_g$ we have a natural one-to-one correspondence between those observables in classical mechanics which depend upon the coordinates, but not the velocities, and certain quantum-mechanical observables. It turns out *not* to be possible to extend this correspondence to one between all classical-mechanical observables and all quantum-mechanical observables. Indeed since classical mechanics is only a limiting case of quantum mechanics there is no reason to expect such a correspondence. Two different quantum-mechanical observables can well coincide in the classical limit. On the other hand there are, as it turns out, well defined quantum analogues for certain of the more important velocity dependent classical-mechanical observables. For example, the formal differential operator $(\hbar/im_j)\partial/\partial x_j$ defines† an observable whose mean value coincides with the time derivative of the mean value of A_{x_j} in every state for which these things have a meaning. In this sense it is the analogue of the classical-mechanical observable known as the x component of the velocity of the j-th particle. To remind ourselves of the analogy we give it the same name in quantum mechanics and do likewise for $(\hbar/im_j)\partial/\partial y_j$ and $(\hbar/im_j)\partial/\partial z_j$. The analogy may be pursued further. The observable corresponding to

$$\sum_{j=1}^{n} m_j\left(\frac{1}{i}\,\frac{\hbar}{m_j}\,\frac{\partial}{\partial x_j}\right) = \frac{\hbar}{i}\sum_{j=1}^{n}\frac{\partial}{\partial x_j}$$

is an integral just as is the corresponding sum in classical mechanics. For obvious reasons it is called the x component of the total momentum and the observable corresponding to the term $(\hbar/i)\partial/\partial x_j$ is called the x component of the momentum of the j-th particle. The y and z components of momentum are defined analogously. In a similar manner one is led to define angular momentum observables and to the discovery that the total angular momentum about any axis is an integral. The angular momentum observables are interesting in that they all have pure point spectra; the spectrum in each case being the set of all integral multiples of $\hbar$. Finally, if one takes the classical energy integral $\sum_{j=1}^{n} \frac{1}{2}m_j(\dot{x}_j^2 + \dot{y}_j^2 + \dot{z}_j^2) + V$ and tries to construct a corresponding quantum-mechanical observable by substituting the analogues already discovered for the $\dot{x}_j$, $\dot{y}_j$, $\dot{z}_j$ and V, one is led to the formal differential operator

$$f \rightarrow -\frac{\hbar^2}{2}\sum_{j=1}^{n}\frac{1}{m_j}\left(\frac{\partial^2 f}{\partial x_j^2} + \frac{\partial^2 f}{\partial y_j^2} + \frac{\partial^2 f}{\partial z_j^2}\right) + Vf$$

which is just $\hbar$ times the differential operator D with which we started. For this reason the observable $\hbar H_D$ is called the energy observable. As we have already remarked it is obviously an integral. The fact that one has natural analogues in quantum mechanics for the energy and momentum observables

† The correspondence between formal differential operators and self adjoint operators on Hilbert space is rather complicated in general. However, for the members of a wide class of first-order operators, including those with which we shall deal, there is a canonical way of passing to a corresponding self-adjoint operator.

lies deeper than the above formal considerations indicate. It can be traced back to the relationship of the observables in question to certain one parameter groups of automorphisms of the system.

12. The quantum mechanics of the atom. The quantum-mechanical model of an atom is not precisely what one would expect from the considerations of the preceding paragraph and the classical picture of an atom as a nucleus surrounded by n electrons. In order to produce results in reasonable agreement with experiment the model resulting from this picture must be modified in three respects. (We simplify the discussion by replacing the (relatively heavy) nucleus by a fixed force field so that an n electron atom leads to an n-body problem.) First $\mathfrak{L}^2(R^{3n})$ must be replaced by the Hilbert space of all square summable functions from R^{3n} to a 2^n-dimensional Hilbert space. Second this new Hilbert space must be replaced by the subspace consisting of those functions f such that $f(T_{ij}(p)) = -f(p)$ for all p in R^{3n} and all i and j with $i \neq j$. Here T_{ij} is the map of R^{3n} onto itself which takes each point into the point obtained by interchanging the coordinates of i-th and j-th particles. Thirdly the restriction of H_D to this subspace must be changed by adding a "small perturbation" which we shall not describe explicitly. The first change has the effect of adding certain new observables—the so-called spin angular moments. The second change eliminates a good many observables—all of those in which the n electrons are not treated symmetrically. The fact that there are observables corresponding to $x_1+x_2+\cdots+x_n$, $x_1^2+x_2^2+\cdots+x_n^2$ *etc.*, but none corresponding to x_1, $x_2, \cdots, x_n$ individually, is interpreted as indicating a new sense in which an electron is not a classical particle. An electron is supposed to have no individuality. It is supposed to be meaningless to make a distinction between a pair of electrons and the same pair with the electrons interchanged. The situation is analogous to that of a pair of identical kinks traveling along in a stretched string. It is possible to have motion in which the kinks seem to have changed places although in fact the state of the string is just as before.

With the indicated modifications, the quantum-mechanical refinement of the classical picture of the atom leads to a model which explains atomic phenomena, including chemistry, in a truly remarkable manner.†

† *Added in proof.* Since this article was sent to the printer, Professor Gleason has improved the result referred to in the footnote on page 51. A proof that there are no measures on the questions other than those associated with states will appear in the Journal of Mathematics and Mechanics.

MULTIDERIVATIVES AND MULTI-INTEGRALS†

KARL MENGER, Illinois Institute of Technology

1. Introduction. The derivative of a function f at a place, a, is defined‡ as the limit of $[f(a')-f(a)]/(a'-a)$ as $a' \to a$. Traditionally, this difference quotient is interpreted geometrically by P^*P'/PP^*, that is, the ratio of the height to the base in the right triangle PP^*P' in Figure 1.

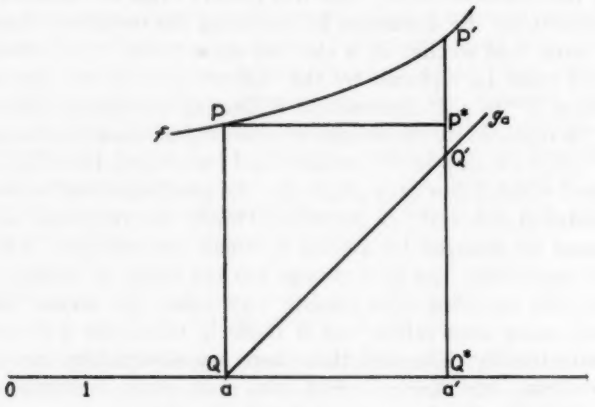


FIG. 1

There is, however, another possible interpretation. Figure 1 shows the line g_a of slope 1 through the point Q below P. Since Q^*Q', the altitude of g_a above Q^*, equals $QQ^*(=PP^*)$, the difference quotient may also be interpreted by P^*P'/Q^*Q', that is, the ratio of the difference between altitudes of f to the difference between altitudes of g_a—altitudes at two neighbor places: a and a'. The line g_a is the graph of a linear function (likewise denoted by g_a) which assumes for any x the value $g_a(x)=x-a$. This linear polynomial serves, as it were, as a gauge function with which one compares f in the neighborhood of a. The derivative of f is obtained by letting a' approach a.

This other interpretation lends itself to far-reaching extensions, recently developed by S. S. Shü and myself. Since our joint note§ is very condensed and rather difficult to read, in Sections 2–5 of the present paper the principal ideas are expounded in detail, stressing motivations and including examples. The last sections contain new material.

† The work on this paper was done as part of Project DA-11-022-ORD-1494 sponsored by the Office of Ordnance Research.

‡ Symbols of functions and function variables are italicized, while numerical variables are printed in roman type.

§ Generalized derivatives and expansions, Proc. Nat. Acad. Sci., vol. 41, 1955, pp. 591–595.

The simplest of the extensions mentioned above results from the comparison of a function f with a *quadratic* gauge function in the neighborhood of *two* places, a and b. More specifically, what serves as the gauge function is the polynomial $g_{a,b}$, assuming for any x the value $g_{a,b}(x) = (x-a)(x-b)$. What will be compared are the difference between chords of f and the difference between chords of $g_{a,b}$—chords between two neighbor pairs of places: between a and b, and between a′ and b′ (see Fig. 2). Any chord is the graph of a linear polynomial. Under certain conditions, also the chord difference quotient (*i.e.*, the ratio of the differences between the chords) is a linear polynomial such that, as a′ and b′ approach a and b, respectively, there exists a linear limit polynomial. For any a and b, this limit polynomial may be called the *bi*derivative of f at a and b. I will here denote it by ${}_2\mathbf{D}f(a, b)$; its value for x by ${}_2\mathbf{D}f(a, b; x)$. For any a and b, the biderivative ${}_2\mathbf{D}f(a, b)$ is a linear function; that is, to say, there exist two numbers $P(a, b)$ and $Q(a, b)$ such that

$$ {}_2\mathbf{D}f(a, b; x) = P(a, b)x + Q(a, b) \text{ for any } x. $$

The numbers $P(a, b)$ and $Q(a, b)$ of course depend on f and, in general in a non-linear way, on a and b.

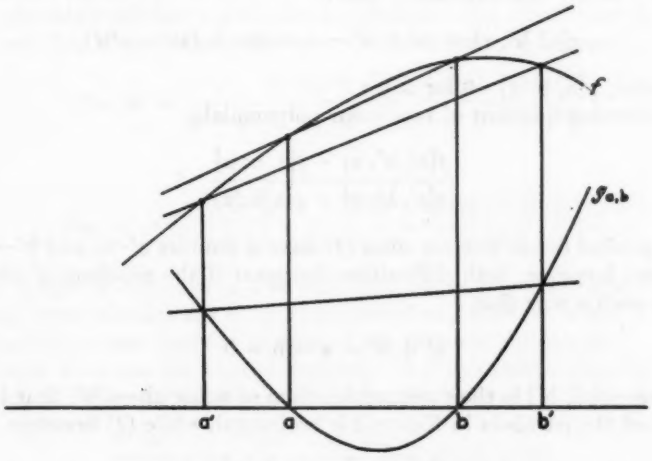


FIG. 2

More generally, ${}_n\mathbf{D}f(a_1, \cdots, a_n)$, the n-derivative of f at n places, results from the comparison of f with the gauge polynomial $g_{a_1,\cdots,a_n}$ of degree n whose value for x is $(x-a_1)\cdot\cdots\cdot(x-a_n)$. The value that this n-derivative assumes for x will be denoted by ${}_n\mathbf{D}f(a_1, \cdots, a_n; x)$. It is a polynomial of degree $n-1$ in x.

The case $n=1$ can be brought in line with this general concept by defining a *uni*derivative ${}_1\mathbf{D}f(a)$ of f at a, namely, as the polynomial of degree 0 (that is,

as the constant function) assuming for any x the value ${}_1Df(a; x) = f'(a)$.

In Sections 2 and 3 it will appear that the prefixed subscript n, introduced in the present paper in denoting the n-derivative, has the advantage of making possible discussions of the operator ${}_nD$ itself and of the result ${}_nDf$ of its application to a function f, instead of confining the study to the polynomials ${}_nDf(a_1, \cdots, a_n)$.

2. The basic definition. The idea outlined in the introduction will now be developed in detail in the case $n = 2$. Extensive use will be made of interpolation polynomials. For any function h and any two numbers c and d ($\neq c$) belonging to the domain of h, the linear polynomial assuming the values $h(c)$ and $h(d)$ for c and d, respectively, will be denoted by $h[c, d]$; its value for x, by $h[c, d; x]$. If c belongs to the domain of h', then $h[c, c]$ will denote the linear polynomial having at c the value $h(c)$ and the derivative $h'(c)$. The graph of $h[c, d]$ is the chord of the curve h between c and d if $d \neq c$, and the tangent to h at c, if $d = c$.

Now let f be a given function and let a and b be two different numbers belonging to the domain of f. The quadratic polynomial called $g_{a,b}$ in the introduction will now, for the sake of brevity, be denoted by g. Thus $g(x) = (x-a)(x-b)$ for any x. For any two numbers a′ and b′

$$g[a', b'; x] = (a' + b' - a - b)x + (ab - a'b'). \tag{1}$$

In particular, $g[a, b; x] = 0$ for any x.

The following quotient of two linear polynomials

$$\frac{f[a', b'; x] - f[a, b; x]}{g[a', b'; x] - g[a, b; x]} \tag{2}$$

is not in general linear in x nor does (2) have a limit as $a' \to a$ and $b' \to b$. Most fortunately, however, both difficulties disappear if the numbers a′ and b′ are chosen in such a way that

$$a' + b' - a - b = 0. \tag{3}$$

In this case, $g[a', b']$ is the constant function of value $ab - a'b'$; that is to say, the chord of the parabola in Figure 2 is horizontal, while (2) becomes

$$\frac{f[a', b'; x] - f[a, b; x]}{ab - a'b'} \tag{4}$$

which is linear in x. In order to verify (under very natural assumptions about f) that (4) possesses a (linear) limit function as $a' \to a$ and $b' \to b$, it is convenient to make use of (3) by setting

$$b' - b = u = a - a'.$$

This makes (4) equal to

$$(5)\qquad \frac{1}{u(b-a+u)}\left\{\frac{f(b+u)-f(a-u)}{b-a+2u}-\frac{f(b)-f(a)}{b-a}\right\}x + \frac{1}{u(b-a+u)}\left\{\frac{(b+u)f(a-u)-(a-u)f(b+u)}{b-a-2u}-\frac{bf(a)-af(b)}{b-a}\right\}.$$

The limit of this quotient as $u\to 0$ can be determined by l'Hospital's Rule under the assumptions that $f'(a)$ and $f'(b)$ exist. (These are the natural assumptions about f alluded to previously.) One readily finds that the limit of (5) as $u\to 0$ is

$$(6)\qquad \frac{1}{(b-a)^2}\left\{f'(a)+f'(b)-2\frac{f(b)-f(a)}{b-a}\right\}x + \frac{1}{(b-a)^2}\left\{(a+b)\frac{f(b)-f(a)}{b-a}-bf'(a)-af'(b)\right\}.$$

For any a and b ($\neq a$), the linear function assuming for any x the value (6) will be called the biderivative of f at (a, b) and will be designated by ${}_2\mathbf{D}f(a, b)$, while (6) will be denoted by ${}_2\mathbf{D}f(a, b; x)$. Clearly, ${}_2\mathbf{D}f(a, b) = {}_2\mathbf{D}f(b, a)$. For any a and b, the two coefficients of the linear expression (6), which have been called $P(a, b)$ and $Q(a, b)$ in the introduction, may be written as follows:

$$P(a, b) = \frac{2}{(b-a)^2}\left\{\frac{f'(a)+f'(b)}{2}-\frac{f(b)-f(a)}{b-a}\right\}$$

and

$$Q(a, b) = \frac{a+b}{(b-a)^2}\left\{\frac{f(b)-f(a)}{b-a}-\frac{af'(b)+bf'(a)}{a+b}\right\}.$$

Thus the graph of ${}_2\mathbf{D}f(a, b)$ is a straight line whose slope is >0 or $=0$ or <0 depending upon whether the slope of the chord of f between a and b is greater than, equal to, or less than the average of the slopes of f at a and b.

The case $b=a$ can be treated similarly. If $f'''(a)$ exists, then by repeated application of l'Hospital's Rule one finds that the limit of (4) as $u\to 0$ is

$$(7)\qquad \frac{f'''(a)}{6}x+\frac{f''(a)}{2}-\frac{af'''(a)}{6} = {}_2\mathbf{D}f(a, a; x).$$

Summarizing one can say:

The symbol	read	denotes
${}_2\mathbf{D}$	the biderivative	an operator
${}_2\mathbf{D}f$	the biderivative of f	a 2-parameter family of linear functions
${}_2\mathbf{D}f(a, b)$	the biderivative of f at (a, b)	a linear function
${}_2\mathbf{D}f(a, b; x)$	the biderivative of f at (a, b) evaluated at x	a number

The last number exists for any function f and any two numbers a and b such that $f'(a)$ and $f'(b)$ exist, if $b \neq a$, and $f'''(a)$ exists, if $b = a$.

More generally, if A is any (unordered) system of n numbers (that is, a finite set of numbers, each endowed with a multiplicity $\geqq 1$, the sum of all multiplicities being equal to n), then the n-derivative of f at A is a polynomial ${}_nDf(A)$ of degree $n-1$. It is defined as the limit of difference quotients

$$\frac{f[A'] - f[A]}{g[A'] - g[A]} \quad \text{as } A' \to A,$$

where the gauge function g is a polynomial of degree n, namely, the polynomial whose value for x is the product of the factors $x-a$ for all a in the system A. In order to make the difference quotient a polynomial of degree $n-1$ and to assure the existence of a limit as $A' \to A$ (under differentiability assumptions concerning f), one must choose A' in such a way that $g[A']$ is a constant polynomial. This means, in generalization of (3), that the sums of the numbers in A' and in A must be equal; but also the sums of the products of the pairs in A' and in A must be equal; and so on, for all elementary symmetric functions except the last, that is, the products of the numbers in A' and in A. Then the denominator of the difference quotient, that is, $g[A'] - g[A]$, is the constant function whose value is the difference between these products.

While the introductory geometrical interpretation of multiderivatives was confined to real numbers and real functions, clearly all algebraic and analytic discussions can be extended verbatim to complex numbers and complex functions.

3. Examples and simple theorems. For the n-th power function, assuming for any x the value x^n, the coefficients in (6), that is, $P(a, b)$ and $Q(a, b)$, can be shown to be polynomials in a and b of degrees $n-3$ and $n-2$, respectively. More specifically, one finds

$f(x)$	${}_2Df(a; b; x)$
x	0
x^2	1
x^3	$x + (a + b)$
x^4	$2(a + b)x + (a^2 + b^2)$
x^5	$(3a^2 + 4ab + 3b^2)x + (a^3 - a^2b - ab^2 + b^3)$
x^6	$(4a^3 + 6a^2b + 6ab^2 + 4b^3)x + (a^4 - 2a^3b - 3a^2b^2 - 2ab^3 + b^4)$
x^n	$\sum_{k=0}^{n-2} (k+1)(n-k-2)a^{n-k-2}b^k x + \sum_{k=0}^{n-2} [(k+1)^2 - kn]a^{n-k-2}b^k.$

These formulae remain valid if $b = a$, in which case they yield

$${}_2Df(a, a; x) = \frac{n(n-1)(n-2)}{6} a^{n-3}x - \frac{n(n-1)(n-5)}{6} a^{n-2},$$

in agreement with (7).

If

$$f(x) = \begin{cases} 0 \text{ for } x < 0 \\ 1 \text{ for } x > 0 \end{cases},$$

then

$${}_2Df(ab; x) = \frac{2}{(b-a)^2}\, x + \frac{a+b}{(b-a)^2}.$$

If $f(x) = x$ and $a < 0 < b$, then

$${}_2Df(a, b; x) = -\frac{2(a+b)}{(b-a)^3}\, x + \frac{2(a^2+b^2)}{(b-a)^3}.$$

The operator ${}_nD$ associates a polynomial of degree $n-1$ with any function and any system of n numbers such that the function has certain differentiability properties at the numbers in the system. Clearly, for any integer n, the operator ${}_nD$ is *linear* in the following sense:

$${}_nD(f_1 + f_2) = {}_nDf_1 + {}_nDf_2 \quad \text{and} \quad {}_nD(cf) = c\, {}_nDf.$$

For instance, in the case $n = 2$ the former formula has the following meaning: If $f = f_1 + f_2$, then for any two numbers a and b for which the biderivatives of f_1 and f_2 exist and for any number x:

$${}_2Df(a, b; x) = {}_2Df_1(a, b; x) + {}_2Df_2(a, b; x).$$

The biderivative of any linear function is the constant function of value 0, wherefore, in the theory of biderivatives, the linear functions play a role similar to that of the constant functions in the classical theory of derivatives (or in the theory of uniderivatives). If l is a linear function, then

$${}_2D(f + l) = {}_2Df \quad \text{for any } f.$$

On the other hand, as the reader can easily verify, ${}_2D(l \cdot f)$ is not in general equal to $l \cdot {}_2Df$. There is, however, a generalization to multiderivatives of the following classical statement: If $f(x) = h(x) \cdot (x-a)$ for every x, then $f'(a) = h(a)$. The extension to biderivatives reads:

If $f(x) = h(x)(x-a)(x-b)$ for every x, then ${}_2Df(a, b) = h[a, b]$; that is to say, ${}_2Df(a, b; x) = h[a, b; x]$ for every x.

4. Multiderivatives of higher order. In our joint paper, Shü and I have introduced multiderivatives of second and higher order by recursive definitions. Higher order multiderivatives can also be introduced by direct limiting processes just as $f''(a)$ can be defined as the limit of

$$\frac{f(a') - f(a) - f'(a) \cdot (a' - a)}{(a' - a)^2/2!} \quad \text{as} \quad a' \to a.$$

For instance, the second order biderivative of f at (a, b) may be defined as the limit of

$$\frac{f[a', b'] - f[a, b] - {}_2Df(a, b)\cdot g[a', b']}{g^2[a', b']/2!}$$

as (a', b') approaches (a, b) in such a way that $b'-b=a-a'=u$ and, therefore, $g[a', b']=ab-a'b'=u(b-a+u)$. If $b\neq a$, the repeated application of l'Hospital's Rule yields

$${}_2D^2f(a, b; x) = \left\{\frac{f''(b) - f''(a)}{(b-a)^3} - 6\,\frac{f'(b)+f'(a)}{(b-a)^4} + 12\,\frac{f(b)-f(a)}{(b-a}\right\} x$$
$$+ \left\{\frac{bf''(a) - af''(b)}{(b-a)^3} + 2\,\frac{(a+2b)f'(a)+(b+2a)f'(b)}{(b-a)^4}\right.$$
$$\left. - 6\,\frac{(a+b)(f(b)-f(a))}{(b-a)^5}\right\}$$

provided that $f''(a)$ and $f''(b)$ exist. For instance, the second order biderivative of all polynomials of a degree $\leqq 3$ is the constant function of value 0.

$f(x)$	$\frac{1}{2}{}_2D^2f(a, b; x)$
x^2	0
x^4	1
x^5	$x+2a+2b$
x^6	$(3a+3b)x+3a^2+3ab+3b^2$
x^7	$(6a^2+9ab+6b^2)x+4a^3+3a^2b+3ab^2+4b$
x^8	$(10a^3+18a^2b+18ab^2+10b^3)x+5a^4+2a^3b+2ab^3+5b^4$
x^n	$\sum_{k=0}^{n-3}\frac{(k+2)(k+1)}{2}\cdot\frac{(n-k-3)(n-k-4)}{2}a^{n-k-4}b^kx+\sum_{k=0}^{n-4}[(k+2)^2-kn]\frac{(k+1)(n-k-3)}{2}a^{n-k-4}b^k$

The preceding expression for $\frac{1}{2}{}_2D^2f(a, b; x)$ has been derived by induction. At my suggestion, Dr. Berthold Schweizer has computed the m-th order derivative of the n-th power function. If $f(x)=x^n$ and one sets

$$\frac{1}{m!}\,{}_2D^m f(a, b; x) = P_n^{(m)}(a, b)x + Q_n^{(m)}(a, b),$$

then Schweizer found

$$P_n^{(m)}(a, b) = \sum_{k=0}^{n-2m-1}\binom{k+m}{m}\binom{n-m-k-1}{m}a^{n-2m-k-1}b^k$$

and

$$Q_n^{(m)}(a, b) = P_{n+1}^{(m)}(a, b) - (a+b)\cdot P_n^{(m)}(a, b).$$

The principal application of $f^{(n)}(a)$ (that is, the classical n-th order derivative of f at a) is to expansions of $f(x)$ in powers of $x-a$. For, according to Taylor,

$$f(x) = f(a) + f'(a)\cdot(x-a) + \tfrac{1}{2}f''(a)\cdot(x-a)^2 + \cdots .$$

In the realm of complex functions, Cauchy expanded $f(z)$ in powers of $z-a$ with coefficients like those in Taylor's expansion; that is to say, the coefficient of $(z-a)^m$ in Cauchy's series is $(1/m!)f^{(m)}(a)$. Moreover, Cauchy identified $f^{(m)}(a)$ and, therefore, that coefficient of $(z-a)^m$ with a line integral along any rectifiable path located within the domain of analyticity of f and encircling the point, a, once in the counterclockwise sense.

Just 100 years ago, Jacobi initiated the study of the expansion of $f(z)$ in powers of $(z-a)(z-b)$ and, more generally, in powers of a polynomial p of higher degree and evaluated the coefficients.† Just 50 years ago, Kienast‡ identified the coefficients with line integrals extending those of Cauchy. But only the new concept of multiderivatives of any order yields expansions with coefficients like those in Taylor's series. Setting $f[a, b] = {}_2D^0f(a, b)$, Shü and I have proved:

$$f(z) = {}_2D^0f(a, b; z) + {}_2Df(a, b; z)(z-a)(z-b) + \tfrac{1}{2}{}_2D^2f(a, b; z)[(z-a)(z-b)]^2 + \cdots .$$

More generally, in expanding f in powers of a polynomial p of degree n, we identified the coefficient of the m-th power of p with the n-derivative of m-th order of f at the system A_p of the zeros of p divided by m!

$$f(z) = {}_nD^0f(A_p; z) + {}_nDf(A_p; z)\cdot p(z) + \tfrac{1}{2}{}_nD^2f(A_p; z)\cdot p^2(z) + \cdots .$$

5. Multiderivatives with respect to general gauge functions. In Section 2, the biderivative of f at (a, b) has been defined as the limit of the quotients (2) as $(a', b')\to(a, b)$ subject to the condition (3). The denominator in (2) is the difference of interpolation polynomials for the function g, which at x assumes the value $(x-a)(x-b)$. This gauge function g is, as it were, adjusted to the pair (a, b) at which the biderivative of f is to be defined. Its graph is the parabola with zeros at a and b. But the same denominator in (2) would result if g were replaced by the function j^2 assuming for any x the value $j^2(x) = x^2$. Indeed,

$$j^2[a', b'] - j^2[a, b] = g[a', b'] - g[a, b].$$

---

† Crelle's Journal, vol. 53, 1856, pp. 103–126. *Cf.* also C. G. J. Jacobi, Gesammelte Werke, vol. 6, pp. 203–230.

‡ A. Kienast, Inaugural Dissertation, University of Zürich, 1906. *Cf.* also P. Montel, Leçons sur les Séries de polynomes à une Variable Complexe, Paris, 1910, pp. 45–54.

B. Schweizer obtained the expressions for $P_n^{(m)}(a, b)$ and $Q_n^{(m)}(a, b)$ by using Kienast's form of the coefficients in the expansion of the n-th power function in powers of the polynomial assuming for any x the value $(x-a)(x-b)$. This computation of the m-th order biderivative is based on the expansion theorem given *l.c.*§ p. 58.

Yet j^2 is in no way adjusted to the particular pair (a, b), wherefore j^2 may serve as the gauge function in defining the biderivative of f at any pair (a, b).

This situation suggested to Shü and myself the possibility of an important generalization. For any differentiable function g, we define a biderivative of f with respect to g, namely, as the limit of the difference quotients (2) as $(a', b') \to (a, b)$ subject to the condition

$$g[a', b'; x] - g[a, b; x] = \text{const.}$$

In view of the last restriction, a linear limit function exists. I will here denote its value for x by ${}_2D_g f(a, b; x)$; the linear function itself, by ${}_2D_g f(a, b)$; the two-parameter family of linear functions in this way associated with f, by ${}_2D_g f$; and the operator, by ${}_2D_g$.

The operator ${}_2D$, discussed in Section 2, clearly is the special case obtained if $g = j^2$. In a formula, ${}_2D = {}_2D_{j^2}$, just as the ordinary uniderivative ${}_1D$ is the uniderivative relative to the identity function j assuming for any x the value $j(x) = x$.

In our joint paper, Shü and I have shown that the m-th order biderivatives of f with respect to g are essentially the coefficients of the m-th power of g in an expansion of f in powers of g. Such expansions considerably generalize those of Jacobi, which are in powers of a polynomial.

6. The most general biderivative of f at (a, b). In the preceding definition of ${}_2D_g f$, one function g has served as the gauge function for any pair of numbers (a, b). The present section is devoted to the most general biderivative of f at one particular pair (a, b). In Section 2, a gauge parabola has been used, and ${}_2Df(a, b; x)$ has been defined as the limit of the expression (5) as $u \to 0$. The limit, obtained by l'Hospital's Rule, remains unchanged if, in the denominator of (5), $u(b-a+u)$ is replaced by $(b-a)u$. The denominator $u(b-a+u)$ is the altitude at $a-u$ and $b+u$ of the gauge parabola, whereas $(b-a)u$ is the altitude at these two places of the V-shaped polygon consisting of the tangents to the parabola at a and at b, respectively. This polygon (which has vertical symmetry) is the graph of the function assuming the value

$$(a - b)(x - a) \text{ for any } x \leqq \tfrac{1}{2}(a + b),$$
$$(b - a)(x - b) \text{ for any } x \geqq \tfrac{1}{2}(a + b).$$

This situation suggests the definition of a more general biderivative of f at (a, b) by comparing f with any V-shaped polygon, namely, for any two unequal numbers p and q, with the polygon $v_{p,q}$ having the altitude

$$p(x - a) \text{ for any } x \leqq \frac{bq - ap}{q - p},$$
$$q(x - b) \text{ for any } x \geqq \frac{q - p}{bq - ap}.$$

The polygon $v_{p,q}$ (which is vertically skew, unless $p=-q$) has the altitude t for $b+t/q$ and $a+t/p$. Setting $1/q-1/p=r$, one has to replace (5) by

$$\frac{1}{t}\left[\frac{f(b+t/q)-f(a+t/p)}{b-a+rt}-\frac{f(b)-f(a)}{b-a}\right]x + \frac{1}{t}\left[\frac{(b+t/q)f(a+t/p)-(a+t/p)f(b+t/q)}{b-a+rt}-\frac{bf(a)-af(b)}{b-a}\right].$$

Its limit as $t\to 0$ is the biderivative of f with respect to $v_{p,q}$, which assumes at x the following value:

$$\frac{1}{(b-a)}\left\{\left[\frac{f'(b)}{q}-\frac{f'(a)}{p}-\left(\frac{1}{q}-\frac{1}{p}\right)\frac{f(b)-f(a)}{b-a}\right]x + \left[\left(\frac{b}{p}f'(a)\right)-\frac{a}{q}f'(b)\right)+\left(\frac{a}{q}-\frac{b}{p}\right)\frac{f'(b)-f'(a)}{b-a}\right]\right\}.$$

It is easily seen that by varying p and q one can obtain the biderivative at (a, b) of f with regard to any gauge function g. All one has to do is to set

$$p=g'(a)-\frac{g(b)-g(a)}{b-a} \quad \text{and} \quad q=g'(b)-\frac{g(b)-g(a)}{b-a}.$$

7. A geometric representation of the biderivative. The uniderivative, ${}_1Df$, is a one-parameter family of constant functions (see Section 1). Geometrically, the function corresponding to a particular value, a, of the parameter may be represented by the horizontal line of altitude $f'(a)$ carrying, as it were, a bead at the abscissa, a, connecting the line visibly with the corresponding value of the parameter. The graph of the classical derivative function, f', is, as it were, a string through all those beads. Knowing this string f' one can, of course, reconstruct the uniderivative ${}_1Df$ by attaching a horizontal line to each bead.

The biderivative ${}_2Df$ is a two-parameter family of linear functions. But the representation by a two-parameter family of straight lines, each carrying two beads, would not be very transparent. The question arises whether the biderivative is not capable of a simpler geometric representation. At first, one might hope to find a curve through all the beads such that each straight line of the family is a chord of that curve joining two beads. Unfortunately, however, except in trivial cases such a curve does not exist.

The chords of any curve h form a two-parameter family. The altitude at x of the line corresponding to the parameters (a, b) is

$$M(a, b)x + N(a, b),$$

where

$$M(a, b)=\frac{h(b)-h(a)}{b-a} \quad \text{and} \quad N(a, b)=\frac{bh(a)-ah(b)}{b-a}.$$

The functions M and N satisfy the following conditions:

1) They are symmetric; that is to say, $M(a, b) = M(b, a)$ and $N(a, b) = N(b, a)$;
2) $a(\partial M/\partial a)(a, b) + (\partial N/\partial b)(a, b) = 0$ and $b(\partial M/\partial b)(a, b) + (\partial N/\partial a)(a, b) = 0$.

In view of the symmetry of M and N, the two conditions 2) are equivalent.

Conversely, if two functions M and N satisfy conditions 1) and 2), then there exists a curve whose chord family is described by M and N. However, if ${}_2\mathbf{D}f(a, b; x) = P(a, b)x + Q(a, b)$, then one readily shows that, except in the case where f is a quadratic function, P and Q do not satisfy condition 2). Incidentally, this fact also proves the insolubility of bidifferential equations of the form

$$ {}_2\mathbf{D}f = {}_2\mathbf{D}^0 f;$$

that is to say,

$$ {}_2\mathbf{D}f(a, b) = f[a, b] \quad \text{or} \quad {}_2\mathbf{D}f(a, b; x) = f[a, b; x]$$

for any a, b, and x.

Of course, one might represent the biderivative ${}_2\mathbf{D}f$ by two surfaces, the graphs of the functions P and Q. But there exists a more interesting interpretation by a single surface, which I will denote by ${}_2f'$. It is the surface having, at any (a, b), the height

$$ {}_2f'(a, b) = {}_2\mathbf{D}f(a, b; a).$$

This single surface obviously permits the reconstruction of the two-parameter family of straight lines ${}_2\mathbf{D}f$. All one has to do in order to construct the line corresponding to the parameter values (a, b) is to join in the cartesian plane the points

$$ (a, {}_2f'(a, b)) \quad \text{and} \quad (b, {}_2f'(b, a)).$$

In particular, for the power functions one finds:

$f(x)$	${}_2f'(a, b)$
x	0
x^2	1
x^3	$2a + b$
x^4	$3a^2 + 2ab + b^2$
. . .	. . . . .
x^n	$\sum_{k=0}^{n-2} (n - k - 1)a^{n-k-2}b^k$

If $f''(a)$ exists, one can easily show that

$$ \lim_{b \to a} {}_2f'(a, b) = f''(a),$$

which, according to (7) equals ${}_2\mathbf{D}f(a, a; a)$ if ${}_2\mathbf{D}f(a, a)$ exists. For instance,

if $f(x)=x^n$, then ${}_2f'(a, a)=\frac{1}{2}n(n-1)a^{n-2}$;
if $f(x)=|x|$, then ${}_2f'(a, b)=2b/(b-a)^2$, which of course has no limit as $b\to a$.

Clearly, also ${}_2\mathbf{D}^n f$ can be represented by a two-place-function or surface ${}_2f^{(n)}$. Similarly, n-derivatives can be represented by n-place-functions.

8. Multi-integration. The classical integration process associates with any continuous function g a two-place function assuming for any pair of numbers, a, a′, the value $\int_a^{a'} g(t)dt$—briefly, $\int_a^{a'} g$. If, in particular, $g=f'$, then $\int_a^{a'} g = f(a')-f(a)$.

In order to make it possible for the integrand to be a uniderivative (rather than a classical derivative function), one might replace g by a one-parameter family of constant functions. The question then arises whether there are extensions of this concept in the direction of multi-integration.

The situation will be illustrated in the case of the bi-integral of a given one-parameter family of linear functions. Let

$$M(u, v)x + N(u, v)$$

be the value for x of the linear function corresponding to the parameter pair (u, v). Can one define a process associating with this family of linear functions a 4-place function whose value for the two pairs (a, b) and (a′, b′) deserves the name of a bi-integral—more precisely, the name of the value of the bi-integral of the family from (a, b) to (a′, b′)? A proper symbol for that value would be

$${}_2\int_{(a,b)}^{(a',b')} (Mx + N).$$

In particular, does $M(u, v)x+N(u, v)={}_2\mathbf{D}f(u, v; x)$ imply

$${}_2\int_{(a,b)}^{(a',b')} (Mx + N) = f[a', b'] - f[a, b]?$$

Such a process indeed exists provided that $a'+b'=a+b$. Set $u=b'-b=a-a'$; choose a positive integer n, and set up the product sum

$$\sum_{k=1}^{n}\left[M\left(a - k\frac{u}{n},\ b + k\frac{u}{n}\right)x + N\left(a - k\frac{u}{n}b + k\frac{u}{n}\right)\right]\left(b - a + 2k\frac{u}{n}\right)\frac{u}{n}\cdot$$

The limit of this expression as $n\to\infty$ clearly is

$$\int_0^1 [M(a - ut, b + ut)x + N(a - ut, b + ut)](b - a + 2ut)udt.$$

For instance, if $M(v, w)=2(v+w)$ and $N(v, w)=v^2+w^2$ for every v and w, then M and N determine the biderivative of the 4-th power function, f^4; that is to say,

$$M(v, w)x + N(v, w) = {}_2Dj^4(v, w; x) \text{ for any } v, w, \text{ and } x;$$

and, correspondingly,

$${}_2\int_{(a,b)}^{(a-u,b+u)} (Mx + N) = j^4[a - u, b + u] - \frac{j^4}{4}[a, b] \text{ for every } a, b, \text{ and } u.$$

The addition of any linear function to j would, of course, not change either the biderivative or the difference of the chords. In this respect, linear functions play the same role in bi-integration that constant functions play in uni-integration.

While the reciprocity between biderivation and bi-integration is restricted to what might be called the symmetric case (that is, to a bi-integration from (a, b) to a pair (a−u, b+u)), a linear bi-integration operator exists for the transition from any pair (a, b) to any pair (a′, b′). This operator associates a four-place function with every one-parameter family of linear functions.

GENERALIZED JACOBI EXPANSIONS AND CORRESPONDING DERIVATIVES*

GORDON PALL, Illinois Institute of Technology

1. Introduction. It is natural to ask whether by suitably generalizing the definition of derivative, it may be possible to formulate as successive derivatives the coefficients A_r in a Jacobi expansion

$$f(x) = A_0(x) + A_1(x)P(x) + \cdots A_r(x)P(x)^r + \cdots, \tag{1}$$

where P denotes a polynomial of degree n and the coefficients A_r are polynomials of lower degree. Several years ago the writer noticed a property of this kind. If we write $P(x) = x^n - a_{n-1}x^{n-1} - \cdots - a_0$, we can regard the coefficients A_r (for given f) as functions of the $n+1$ parameters $x, a_0, \cdots, a_{n-1}$; then if f is regular at the zeros of P, $k!A_k$ is the k-th partial derivative of A_0 with respect to a_0, and various relations connect derivatives with respect to $a_1, \cdots, a_{n-1}$. As part of a project involving the writer, Menger and Shü† recently defined a derivative of a function at a system of points in such a way that the coefficients $k!A_k$ are indeed successive derivatives of f at the system of zeros of P. In their definition, the interpolation polynomial of f at a one-parameter family of n-tuples is used somewhat as in our Section 3, and their derivatives are defined as polynomials whose coefficients are functions, and are accordingly two-place functions, or even $(n+1)$-place functions if the n points are regarded as arbitrary, rather than one-place functions like the elementary derivative.

In the present article the Menger-Shü derivative is somewhat generalized, and is then reconstituted as a one-place function under fairly general conditions. It becomes then remarkably similar to the elementary derivative.

A method of obtaining a large variety of expansions, both new and old, is described at the end of Section 3.

Properties of the generalized derivative are briefly developed in Section 4, and a useful expression analogous to a divided difference formula is found. Analogues of Rolle's and mean-value theorems are discussed in Section 5. Applications to the solution of certain types of differential-functional equations are indicated in Section 6. As an interesting example, but independently of the general theory, a very simple way of developing the theory of Jacobi series is given in Section 2. The simplicity is induced here, first by separating from f the discussion of the mapping associated with P, and then reducing the study to that of certain coefficient functions rather than of f itself. The same idea lies behind the general algorithm of Section 3.

The author wishes to express his gratitude to both Professors Menger and Shü for their inspiration and help in connection with the project and this paper. References to the literature will be found in their work.†

* This work was sponsored by the Office of Ordnance Research under Project TB 2-0001(888).

† K. Menger and S. S. Shü, Generalized derivatives and expansions, Proc. Nat. Acad. Sci., vol. 41, 1955, pp. 591–595. Further references to the literature will be found herein.

2. An example: expansion in powers of a polynomial. It will be instructive to precede the general theory of Section 3 with an interesting special case. The treatment is independent of the general theory and really simple.

Consider the mapping between two complex planes defined by

$$w = z^2 - 2az + b, \tag{1}$$

where a and b are given complex numbers. Set $c = b - a^2$. If w_0 is any point different from c in the w-plane, the interior W of the circle with center w_0 and passing through c is mapped into the two half interiors Z_1 and Z_2 of a lemniscate with vertex at the point a in the z-plane.

Equation (1) has two solution-functions $z_1 = \phi_1(w)$ and $z_2 = \phi_2(w)$ which are regular in W and map it one-to-one onto Z_1 and Z_2 respectively.

Consider now a function f regular in both Z_1 and Z_2. We allow the possibility that, in the sense of analytic continuation, f may be a different analytic function in the two parts. With f are associated two functions a_0 and a_1 defined on W as follows:

$$f(z) = a_0(w) + a_1(w)z \tag{2}$$

under $z = z_1 = \phi_1(w)$ and $z = z_2 = \phi_2(w)$, w in W.

On solving for a_0 and a_1, one finds

$$a_0(w) = \frac{z_1 f(z_2) - z_2 f(z_1)}{z_1 - z_2}, \qquad a_1(w) = \frac{f(z_1) - f(z_2)}{z_1 - z_2}. \tag{3}$$

Since f is regular in Z_1 and Z_2, and since $\phi_1(w) \neq \phi_2(w)$ in W, it is evident from (3) that a_0 and a_1 are regular in W. Hence we can expand

$$a_j(w) = \sum_{n=0}^{\infty} c_{nj}(w - w_0)^n \qquad (j = 0, 1), \tag{4}$$

in series convergent in W. This gives at once the expansion

$$f(z) = \sum_{n=0}^{\infty} (c_{n0} + zc_{n1})(h(z) - h(z_0))^n, \tag{5}$$

valid in Z_1 and Z_2. Here $h(z)=z^2-2az+b$ and $h(z_0)=w_0$.

Consider now the expression

$$a_0'(w)+za_1'(w) \tag{6}$$

obtained from (2) by differentiating the coefficient functions. The one-place function f^* defined by

$$f^*(z)=a_0'(h(z))+za_1'(h(z)) \tag{7}$$

bears the same relation to (6) (under (1)) as f does to the right member of (2). We will call the function f^* the h-derivative of f, and designate it by D_hf. It has in Z_1 and Z_2 the expansion

$$\sum_{n=1}^{\infty} n(c_{n0}+zc_{n1})(h(z)-h(z_0))^{n-1}. \tag{8}$$

We may call the linear function $c_{00}+zc_{01}$ the *linear value of f at z_0*, or perhaps better, *at z_0 and $2a-z_0$*. (Here z_0 and $2a-z_0$ are *conjugate* points, for which h has the same value.) It follows that the linear function $c_{10}+zc_{11}$ is the linear value of D_hf; and that $n!(c_{n0}+zc_{n1})$ is the linear value at z_0 and $2a-z_0$ of D_h^nf, where

$$D_h^nf=D_h(D_h^{n-1}f) \qquad (n=2,3,\cdots).$$

If the two parts of f are not analytic continuations of one another, then c must be a singularity (and, if a path around c is possible, a branch point) of a_0 or a_1. If f is given as regular in a neighborhood of z_0 and $2a-z_0$, smaller than Z_1 and Z_2, a circle of some radius less than $|c-w_0|$ can be used, and an expansion of f obtained in a corresponding region.

The ideas of this section extend in an obvious manner to the expansion of analytic functions in powers of given polynomials of any degree.

3. The interpolation value and bifunction of a function. Let $W, Z_1, \cdots, Z_n$ be subsets of a ring R. Let W be mapped one-to-one onto Z_i by a function p_i $(i=1,\cdots,n)$. We may denote a sequence of corresponding elements by $w, z_1, \cdots, z_n$, so that $z_i=p_i(w)$. We shall, unless otherwise indicated, assume that for every given w the n elements $z_1, \cdots, z_n$ are distinct.

Write Z for the set of all elements in $Z_1, \cdots, Z_n$. By an *interpolation basis* is meant a sequence of n functions $q_1, \cdots, q_n$ defined on Z to R, such that for every function f on Z to R there exist for every given w uniquely determined elements $a_1(w), \cdots, a_n(w)$ of R such that

$$f(z_i)=a_1(w)\cdot q_1(z_i)+\cdots+a_n(w)\cdot q_n(z_i) \qquad (i=1,\cdots,n). \tag{1}$$

In this manner n coefficient functions $a_1, \cdots, a_n$ on W to R are associated with f.

A two-place function of the form $b_1q_1+\cdots+b_nq_n$, where b_i are functions

"of w", and the q_i "of z", will be called a *bifunction*. The bifunction $a_1q_1+\cdots+a_nq_n$ defined by (1) will be termed the *interpolation function of* f relative to the basis. For a given value of w, the function "of z"

$$a_1(w)\cdot q_1+\cdots+a_n(w)\cdot q_n \tag{2}$$

will be called the *interpolation value* at w, or the w-*value* of f. Such w-values will have derivative zero in the sense to be defined in Section 4, and will play the role of constants.

Examples easily show that, in general, it is not true conversely that every bifunction is the interpolation function of some function g on Z to R. It is natural to ask whether the p's and q's can be selected so that every bifunction will be an interpolation function. We shall find a necessary and sufficient condition for this to be the case, assuming however that q_1 is the constant function 1, and that $z_1,\cdots,z_n$ are distinct for each w. This condition is that there exists a function h on Z to W such that

$$h(p_i(w))=w \qquad (w \text{ in } W;\ i=1,\cdots n). \tag{3}$$

In other words the inverse mapping is also one-valued; if the Z's overlap, common points come from the same w. A function h satisfying (3) will be called a *gauge function*.

The necessity of the existence of a gauge function follows from the fact that $w\cdot 1+0\cdot q_2+\cdots+0\cdot q_n$ must be the interpolation function of some function, and if we denote this function by h we have (3). The sufficiency follows from the observation that the one-place function obtained from the bifunction $b_1q_1+\cdots+b_nq_n$ by substituting h for w, that is, the function defined by

$$b_1(h(z))\cdot q_1(z)+\cdots+b_n(h(z))\cdot q_n(z) \qquad (z \text{ in } Z), \tag{4}$$

has the bifunction $b_1q_1+\cdots+b_nq_n$ as its interpolation function. It should be remarked in passing that while this argument is incomplete when equalities are allowed among $z_1,\cdots,z_n$, it can be extended so that the construction in (4) applies to certain other cases, with a suitable modification of the definition of interpolation function.

We thus have a remarkable situation in which there is a unique association between one-place functions f on Z to R and bifunctions. Any linear operator acting on the coefficient functions $b_1,\cdots,b_n$ induces a corresponding linear operator on the one-place functions f; and conversely. For example, if in appropriate cases $b_1,\cdots,b_n$ have a derivative at w, then there is defined a corresponding "derivative" of f at z (where z may be identified with any of $z_1,\cdots,z_n$). Again, if the coefficient functions can be expanded in a series of any kind (for example, a power series or trigonometric series), or can be represented by certain types of integrals, then there is a corresponding expansion or integral representation of f; and conversely. To establish the validity of such expressions it is only necessary to determine those properties of f which cause the coefficient functions to possess the corresponding expansions. For example

in Section 2, the analyticity of the coefficient functions at points w for which the associated z's are distinct followed from the analyticity at the points z of f.

In cases of equality among $z_1, \cdots, z_n$ for given w, the n equations (1) are no longer independent. In appropriate cases one may adopt a convention like the familiar one of having not only the values of f and $a_1(w)q_1+\cdots+a_n(w)q_n$ agree on $z_1, \cdots, z_n$, but also their derivatives up to order $k-1$ at points of multiplicity k. Menger and Shü have given conditions under which the coefficient functions have certain continuity or differentiability properties, when $q_i(z)=z^{i-1}$.

4. Definition, evaluation and properties of a generalized derivative. We now assume that R is the field of real or complex numbers, and that the sets W, $Z_1, \cdots, Z_n$ are either simply-connected open regions, or open intervals or arcs. We assume $q_1, \cdots, q_n$ to be continuously differentiable on Z up to order n. By definition of interpolation basis, if $z_1, \cdots, z_n$ are distinct for a given w, the determinant $\Delta=|q_j(z_i)|$ cannot vanish.

If the mapping functions $p_1, \cdots, p_n$ are continuous at a point w of W, then the application of Cramer's rule to (1) of Section 3, with z_i replaced by $p_i(w)$ $(i=1, \cdots, n)$, expresses $a_j(w)$ as a quotient of two determinants each of which is continuous in w if f is continuous at each of $z_1, \cdots, z_n$. Hence if f is continuous, the coefficient functions a_j are continuous at points w for which $z_1, \cdots, z_n$ are distinct. Also at such points, if the mapping functions $p_1, \cdots, p_n$ are differentiable up to order k and f is differentiable up to order k, and the q's are likewise differentiable, then the coefficient functions are differentiable up to order k. When equalities occur among $z_1, \cdots, z_n$, the situation can be more complicated, and it is usually necessary to assume continuity or differentiability of $p_1, \cdots, p_n$ of a higher order than that desired for the coefficient functions.

The bifunction defined by

$$a_1'(w)q_1(z)+\cdots+a_n'(w)q_n(z) \qquad (w \text{ in } W,\ z \text{ in } Z), \tag{1}$$

will, when there is a gauge function, be the interpolation function of the one-place function defined for each z of Z by

$$a_1'(h(z))q_1(z)+\cdots+a_n'(h(z))q_n(z). \tag{2}$$

The one-place function so defined will be called the h-derivative of f relative to the system p_i, q_i; or briefly, the *h-derivative of f*, and denoted by $D_h f$.

If f is a function of h, say $\psi(h)$, the corresponding bifunction is $\psi(w)\cdot 1+0\cdot q_2+\cdots$, and hence the h-derivative is $\psi'(h)$. Similarly, the h-derivative of a product in which one of the factors is a function of h is given by a formula analogous to that in the elementary calculus:

$$D_h(f\cdot g(h))=f\cdot g'(h)+D_h f\cdot g(h). \tag{3}$$

If f has the form $c_1q_1+\cdots+c_nq_n$, where the coefficients $c_1, \cdots, c_n$ are constants, then $D_h f=0$. Such functions may be called h-constants. The h-derivative

of a sum of terms is the sum of their h-derivatives. If A is an h-constant, the h-derivative of $A \cdot h^k$ is kAh^{k-1}. (Our discussion here is brief; conditions under which these formulae hold are obvious.)

We will now derive an explicit formula for $D_h f$ in the case where $z_1, \cdots, z_n$ are distinct and f is differentiable at the points $z_1, \cdots, z_n$, and where $h'(z)$ does not vanish at these n points. The expression obtained will be in terms of $z_1, \cdots, z_n$, but becomes a function "of z" by identifying z with any one of these n quantities, and then through the mapping process expressing the others in terms of that one.

From (1) of Section 3 we have, for every w in W,

$$f(p_i(w)) = a_1(w)q_1(p_i(w)) + \cdots + a_n(w)q_n(p_i(w)) \qquad (i = 1, \cdots, n). \tag{4}$$

Differentiating with respect to w, and noting that $h(p_i(w)) = w$ gives $h'(p_i) = 1/p_i'$, we have (assuming existence of all derivatives listed) that

$$a_1' \cdot q_1(p_i) + \cdots + a_n' \cdot q_n(p_i) = (f'(p_i) - \sum_j a_j \cdot q_j'(p_i))/h'(p_i). \tag{5}$$

Hence the left member is the value at $p_i (= p_i(w))$ of

$$(f' - \sum a_j \cdot q_j')/h'. \tag{6}$$

Since the value of (6) at p_i is given by the left member of (5) we can use (6) to evaluate the one-place function $D_h f$. We transform (6) by substituting $a_j(h(z)) = \Delta_j/\Delta$, obtained by applying Cramer's rule to the equations

$$\sum_j a_j(h(z)) \cdot q_j(z_i) = f(z_i) \qquad (i = 1, \cdots, n), \tag{7}$$

where $z_i = p_i(h(z))$. Here Δ is the determinant $|q_j(z_i)|$, and Δ_k is obtained by replacing the column of elements $q_k(z_i)$ by $f(z_i)$, $k = 1, \cdots, n$. Hence by (6), the derivative $D_h f$ has the form

$$\left\{\Delta \cdot f'(z) - \sum_j \Delta_j q_j'(z)\right\} \Big/ (\Delta \cdot h'(z)),$$

and this is seen to be the expansion of

$$\begin{vmatrix} f'(z) & 0 & q_1'(z) & q_2'(z) \cdots q_n'(z) \\ f(z_1) & 1 & q_1(z_1) & q_2(z_1) \cdots q_n(z_1) \\ f(z_2) & 1 & q_1(z_2) & q_2(z_2) \cdots q_n(z_2) \\ \cdot & \cdot & \cdot & \cdot \quad \cdots \quad \cdot \\ f(z_n) & 1 & q_1(z_n) & q_2(z_n) \cdots q_n(z_n) \end{vmatrix} \cdot \frac{1}{\Delta \cdot h'(z)}, \tag{8}$$

where as noted above, z can be identified with any of $z_1, \cdots, z_n$, and the other

conjugates expressed in terms of that one. In the case $q_j(z)=z^{j-1}$, the entire expression (8) is the divided difference of f at the points $z, z_1, \cdots, z_n$, where z coincides with one of the z_i. Presumably the usual conventions for divided differences will extend the validity of this expression for $D_h f$ when equalities occur among the n conjugates of z.

Formula (8) suggests a "rate" interpretation of $D_h f$: it may be described as a weighted measure of the difference between the local rate of change of f at z, and the average rate of change between z and its conjugate points. If $n=2$, the expression (8) reduces to

$$\frac{1}{h'(z)}\left\{f'(z) - q_2'(z)\cdot\frac{f(z^*) - f(z)}{q_2(z^*) - q_2(z)}\right\}, \tag{9}$$

where z^* is the conjugate of z. If $q_2(z)=z$, this becomes

$$\frac{1}{h'(z)}\left\{f'(z) - \frac{f(z^*) - f(z)}{z^* - z}\right\}. \tag{10}$$

If here $h(z)=z^2-rz+s$, then $z^*=r-z$ and $h'(z)=z-z^*$. Graphically this can be described as the one-place function obtained from the graph of $y=f(x)$ in the following manner. Take the difference between the chords at $z+k$ and z^*-k, and at z and z^*, divide by k, take the limit as k tends to zero, thus getting a straight line; divide by $z-z^*$, getting another straight line. Retain of this line only the two points with abscissae z and z^*: the locus of all such point pairs is the graph of $D_h f$.

Notice that the denominator $\Delta\cdot h'(z)$ in (8) is the value of the numerator with f replaced by h.

5. A remark on analogues of Rolle's or mean-value theorems. Consider the case $n=2$, $h(x)=x^2-rx$, r real. On the graph of $y=f(x)$ let the points with abscissae x_1, $a-x_1$, x_2, $a-x_2$ be collinear; in other words, let the linear values of f at x_1 and x_2 be the same. Let f be continuous at the four points and suppose f' exists between x_1 and x_2, and between $r-x_1$, $r-x_2$. By (10) of Section 4, $D_h f$ exists between x_1 and x_2. A direct analogue of Rolle's theorem would require the conclusion that for some point ξ between x_1 and x_2, the linear value of $D_h f$ at ξ should vanish; in other words, the points with abscissae ξ and $a-\xi$ should have a common tangent. Brief examination of a figure will show that in general this is not so, except in the rather special cases in which f is either symmetric or anti-symmetric about the point r; that is, for x between x_1 and x_2, either $f(r-x)=f(x)$, or $f(r-x)=-f(x)$. By writing f as the sum of $f_1(x)=\frac{1}{2}(f(x)+f(a-x))$ and $f_2(x)=\frac{1}{2}(f(x)-f(a-x))$, one can develop (exactly as in the elementary calculus) Taylor expansions with remainder terms. However, these remainder terms involve linear values of the n-th h-derivatives of f_1 and f_2 at different intermediate points.

However, closer examination shows that this type of development, while

interesting because of the analogy with elementary calculus, is really a subterfuge for the expansion of the coefficient functions $a_1(t)$ and $a_2(t)$, where

$$f(x) = a_1(t) + xa_2(t), \text{ under } t = h(x) - h(x_1);$$

and the symmetric and anti-symmetric cases are simply those in which $a_1(t)$ and $a_2(t)$ are linearly dependent. When Z_1 is a real interval, one can construct for the general situation of Sections 3, 4, expansions of the Taylor type with remainder terms involving values at n different intermediate points of derivatives of the coefficient functions. Nonetheless, expansions in which the remainders involve only one intermediate point are also possible in fairly general cases, and indeed, to get such expansions, we need only keep x constant in the expression $a_1(t)\cdot q_1(k) + \cdots + a_n(t)\cdot q_n(k)$, apply Taylor's formula to this function of t, and replace k by x, t by $h(x)$, and the intermediate point τ provides a single intermediate point ξ such that $h(\xi) = \tau$.

6. Certain differential-functional equations. We mention briefly an application to the solution of certain types of differential equations involving values of the solution function at two or more points. If a function h can be found having equal values at such "conjugate points", and if the equation can be expressed in the form $D_h f = g$ for a suitable choice of p's and q's, then the equation can be solved by forming the bifunction of g, integrating the coefficient functions with respect to their argument, and then substituting $h(z)$ for that argument in the resulting expressions.

In some cases, where the conjugate points can be represented in terms of one another by the same analytic expression, one can even avoid using the parameter w, by using $dw = h'(z)dz$. For example, with (10) of Section 5 in mind, let us solve the equation

$$f'(x) - \frac{f(x) - f(1-x)}{2x-1} = x^2.$$

Dividing by $h'(x) = 2x - 1$, we choose functions $a_1(t)$, $a_2(t)$ such that

$$x^2/(2x-1) = a_1(t) + xa_2(t), \qquad (1-x)^2/(1-2x) = a_1(t) + (1-x)a_2(t).$$

Hence $a_2(t) = (2x^2 - 2x + 1)/(2x-1)$, $2a_1(t) + a_2(t) = 1$, and since $dt = (2x-1)dx$,

$$\int a_2(t)dt = \int (2x^2 - 2x + 1)/(2x-1)dx = \tfrac{1}{2}(x^2 - x) + \tfrac{1}{4}\ln|2x-1| + c_1,$$

$$= x^2 - x - 2\int a_1(t)dt - 2c_2,$$

hence

$$f(x) = \int a_1(t)dt + x\int a_2(t)dt$$

$$= (2x^3 - x^2 - x)/4 + ((2x-1)/8)\ln|2x-1| + c_1x + c_2.$$

ON THE SPACE OF ORIENTED LINES IN EUCLIDEAN THREE-SPACE*

TIBOR RADÓ, The Ohio State University

1. Introduction. Throughout this note, the term *line* means *straight line* in Euclidean three-space E_3. An *oriented line* g is a line with one of the two directions on the line designated as the positive direction. Thus an oriented line may be represented geometrically by a line with an attached arrow indicating the positive direction. If g is an oriented line, then $[g]$ will denote the non-oriented line from which g arose. G will denote the set of oriented lines in E_3.

In many important situations arising in analysis and in geometry one has to consider real-valued functions $f(g)$ of oriented lines g, as well as integrals $\int f dm$ of such functions with respect to a properly chosen measure in G. In recent years, several such situations arose in connection with surface area theory (see, for example, E. J. Mickle and T. Radó, *A new geometrical interpretation of the Lebesgue area of a surface*, Duke Math. Journal, vol. 15, 1948, pp. 169–180). The increasing complexity of these situations made it seem desirable to survey the various possibilities in constructing a theory of measure and integration in G, and in particular to study the various possibilities in introducing a topology in G as a basis for the analytical applications. This latter issue led the writer to a number of rather elementary and yet interesting questions which, as experience has shown, can be used as stimulating illustrations in a first course in general topology. The purpose of this note is to present these questions and the corresponding results. The presentation is designed for a reader familiar with general topology who may want to use this material in his own teaching, and hence many simple proofs are omitted.

As regards terminology, the non-oriented line $[g]$ associated above with the oriented line g will be termed the *carrier* of g. Two oriented lines g_1, g_2 are termed *parallel* if their carriers are parallel. If so, g_1 and g_2 may have the same or opposite orientations. In the first case, g_1 and g_2 are said to be *syntactic*, while in the second case they are said to be *antitactic*. If g_1, g_2 are antitactic and $[g_1]=[g_2]$, then g_1, g_2 are termed *opposite*. This terminology has been introduced by the German mathematician Study.

2. Metric spaces. We collect in this section the basic definitions to be used in the sequel. A metric space (X, d) consists of a non-empty set X and a real-valued function $d(x, y)$ of pairs of points (the *distance-function*), such that (i) $0 \leqq d(x, y) < \infty$, (ii) $d(x, y)=0$ if and only if $x=y$, (iii) $d(x, y)=d(y, x)$, (iv) $d(x, y)+d(y, z) \geqq d(x, z)$ (the *triangle inequality*). There arises a natural *limit concept*, to be denoted by $\Lambda(d)$, such that $x_n \to x_0$ if and only if $d(x_0, x_n) \to 0$.

* This study is part of the background material prepared for a research project (on the measure-theoretical foundations of surface area theory) sponsored by the Office of Ordnance Research, U. S. Army, contract number DA 33-019-ORD-2114.

In a metric space (X, d), the family $\Omega(d)$ of open sets is defined as the family of those sets $E \subset X$ which have the following property: for every point $x \in E$ there exists a positive number $\eta = \eta(x, E)$ such that $d(y, x) < \eta$ implies that $y \in E$. On introducing $\Omega(d)$ in this manner, one obtains a topological space (indeed a Hausdorff space). The family $\Omega(d)$ is termed the topology induced in X by the distance-function d. The family $\Gamma(d)$ of closed sets consists of those sets $F \subset X$ whose complements are open. A point x is termed a limit point of a set E if there exists a sequence $x_n \in E$ such that $x_n \to x$. One has then the following direct description of closed sets in (X, d): a set $E \subset X$ is closed if and only if it contains all of its limit points. It follows that if two distance-functions d', d'' yield the same limit concept then they induce the same topology. In symbols: if $\Lambda(d') = \Lambda(d'')$, then $\Omega(d') = \Omega(d'')$. The converse is obviously also true.

Two distance-functions d', d'' in X will be termed *equivalent*, in symbols $d' \sim d''$, if and only if they induce the same topology, that is, if and only if $\Omega(d') = \Omega(d'')$.

Let X be a non-empty set in which a limit concept Λ is assigned in the following sense: certain sequences $\{x_n\}$ of points of X are selected as convergent sequences, and with each convergent sequence $\{x_n\}$ there is associated a unique point $x \in X$ as its limit. Then the aggregate (X, Λ) is termed a limit space. A limit space (X, Λ) is termed metrizable if there exists in X a distance-function d (with the properties (i)–(iv) stated above) such that $\Lambda = \Lambda(d)$. From the facts stated above it follows that if a limit space (X, Λ) is metrizable, then any two distance-functions d', d'' which serve to metrize it yield the same topology in X.

3. Scale-functions. Given a metric space (X, d), a real-valued function $f(t)$, defined for $0 \leqq t < \infty$ and satisfying $0 \leqq f(t) < \infty$, is termed a *scale-function* for (X, d) if $f[d(x, y)]$ is again a distance-function in X (that is, if the function $d'(x, y) = f[d(x, y)]$ satisfies the conditions (i)–(iv) stated in Section 2). Note that we do *not* require that the distance-function $f[d(x, y)]$ should be equivalent to d. If k is a finite positive constant, then clearly kt is a scale-function for (X, d), and indeed this special case accounts for the terminology. The function $f(t)$ defined by $f(0) = 0$, $f(t) = 1$ for $0 < t < \infty$, is a scale-function for (X, d) which yields, in the form $f[d(x, y)]$, the *isolating distance-function* d_i such that $d_i(x, y) = 0$ if $x = y$ and $d_i(x, y) = 1$ if $x \neq y$. The function $t/(1+t)$, $0 \leqq t < \infty$, is a scale-function for (X, d) which yields a bounded distance-function $d_b = d/(1+d)$ equivalent to d. A function $f(t)$ is termed a universal scale-function if it is a scale-function for every metric space (X, d). One verifies readily that the class of universal scale-functions coincides with the class of scale-functions for the Euclidean plane E_2. The constructive determination of all scale-functions is still an open problem. Many remarkable and deep results relating to scale-functions have been obtained, in particular, by Schoenberg and v. Neumann (for details and references, see the excellent monograph by L. M. Blumenthal, *Distance Geometries*, The University of Missouri Studies, vol. 13, 1938).

4. Scale-functions for the Euclidean plane E_2.

4.1. The Euclidean plane E_2 is the metric space whose points $x=(x^1, x^2)$ are ordered pairs of real numbers x^1, x^2, and where the distance of two points $x=(x^1, x^2)$, $y=(y^1, y^2)$ is the Euclidean distance $d_e(x, y)$ given by the formula

$$d_e(x, y) = [(x^1 - y^1)^2 + (x^2 - y^2)^2]^{1/2}.$$

Accordingly, a scale-function $f(t)$ for E_2 is a real-valued function, defined for $0 \leqq t < \infty$, such that

(i) $0 \leqq f(t) < \infty$ for $0 \leqq t < \infty$,

(ii) $f[d_e(x, y)]$ is a distance-function in E_2 (which is *not* required to be equivalent to the Euclidean distance-function d_e).

We proceed to collect some simple facts needed in the sequel.

4.2. A triple (t_1, t_2, t_3) of non-negative, finite real numbers will be termed a *Euclidean triple* if there exist in E_2 three (not necessarily distinct) points x_1, x_2, x_3 such that

$$t_1 = d_e(x_2, x_3), \qquad t_2 = d_e(x_3, x_1), \qquad t_3 = d_e(x_1, x_2).$$

Briefly, (t_1, t_2, t_3) is a Euclidean triple provided that there exists in E_2 a (perhaps degenerate) triangle with side-lengths t_1, t_2, t_3. Obviously, this is the case if and only if t_1, t_2, t_3 satisfy the inequalities

$$t_1 \leqq t_2 + t_3, \qquad t_2 \leqq t_3 + t_1, \qquad t_3 \leqq t_1 + t_2.$$

The following statement is thus obvious.

LEMMA. *A real-valued function $f(t)$, defined for $0 \leqq t < \infty$, is a scale-function for E_2 if and only if the following conditions are satisfied.*

(a) $0 \leqq f(t) < \infty$ *for* $0 \leqq t < \infty$,

(b) $f(0) = 0$, $f(t) > 0$ *if* $t > 0$,

(c) *if (t_1, t_2, t_3) is any Euclidean triple, then $[f(t_1), f(t_2), f(t_3)]$ is also a Euclidean triple.*

4.3. Let $f(t)$ be a scale-function for E_2. Put

$$\lambda = \underset{0<t<\infty}{\text{g.l.b.}} f(t), \qquad l = \liminf_{t \to +0} f(t), \qquad L = \limsup_{t \to +0} f(t).$$

Note that λ, l, L are defined in terms of the behavior of $f(t)$ for positive values of t.

LEMMA 1. *If $0 \leqq t_1 \leqq t_2$, then $f(t_1) \leqq 2f(t_2)$.*

Indeed, (t_1, t_2, t_2) is clearly a Euclidean triple, and thus, by the lemma in 4.2, $[f(t_1), f(t_2), f(t_2)]$ is also a Euclidean triple. Hence $f(t_1) \leqq f(t_2) + f(t_2) = 2f(t_2)$.

LEMMA 2. $\lambda \leqq l \leqq L \leqq 2\lambda$.

LEMMA 3. *If $0 \leqq t_n < \infty$, $n = 1, 2, \cdots$, and $f(t_n) \to 0$, then $t_n \to 0$.*

The proofs of these two lemmas are readily made by using Lemma 1. In the way of a hint, we indicate the verification of the inequality $L \leqq 2\lambda$. Assign $\epsilon > 0$. By the definition of λ, there exists then a t_0 such that

$$0 < t_0 < \infty \text{ and } f(t_0) < \lambda + \epsilon.$$

For $0 < t < t_0$ we have then, by Lemma 1, the inequality $f(t) < 2(\lambda + \epsilon)$, and thus $L \leqq 2(\lambda + \epsilon)$. As $\epsilon > 0$ was arbitrary, it follows that $L \leqq 2\lambda$.

4.4. A *rigid transformation* in E_2 is a mapping $\rho: E_2 \to E_2$ such that

$$d_e[\rho(x), \rho(y)] = d_e(x, y)$$

for every choice of the points x, y in E_2. A distance-function d in E_2 is termed *invariant* if $d[\rho(x), \rho(y)] = d(x, y)$ for every rigid transformation ρ and for every choice of the points x, y in E_2. The following statement is obvious.

LEMMA. *A distance-function d in E_2 is invariant if and only if there exists a scale-function $f(t)$ for E_2 such that*

$$d(x, y) = f[d_e(x, y)]$$

for every choice of the points x, y in E_2.

Thus the determination of all scale-functions for E_2 is equivalent to the determination of all invariant distances in E_2. While this issue is still open, we obtain a readily solvable problem if we ask for all the topologies induced by invariant distance-functions in E_2. Indeed, we have the following

THEOREM. *Let d be an invariant distance-function in E_2. Then either d is equivalent to the Euclidean distance d_e, or else d is equivalent to the isolating distance*

$$d_i(x, y) = \begin{cases} 0 & \text{if } x = y, \\ 1 & \text{if } x \neq y. \end{cases}$$

In other words, the topology $\Omega(d)$ induced by an invariant distance-function d in E_2 is either identical to the Euclidean topology $\Omega(d_e)$ induced by the Euclidean distance d_e or else it is identical to the discrete topology induced by the isolating distance d_i.

Proof. By the lemma above, we have $d(x, y) = f[d_e(x, y)]$, where $f(t)$, $0 \leqq t < \infty$, is a scale-function for E_2. To this $f(t)$ we apply the notations and lemmas of 4.3.

Case 1. $\lambda = 0$. Then we assert that $d \sim d_e$. Indeed, by Lemma 2 in 4.3, the assumption $\lambda = 0$ implies that $l = L = 0$, and thus (since $f(0) = 0$) it follows that $f(t)$ is continuous at $t = 0$. Consider now in E_2 any sequence of points x_n, $n = 0, 1, 2, \cdots$, such that $d_e(x_n, x_0) \to 0$ as $n \to \infty$. By the just established continuity of $f(t)$ at $t = 0$ it follows that

$$d(x_n, x_0) = f[d_e(x_n, x_0)] \to 0.$$

Conversely, let y_n, $n = 0, 1, 2, \cdots$, be any sequence of points in E_2 such that

$d(y_n, y_0) \to 0$ as $n \to \infty$. As $d(y_n, y_0) = f[d_e(y_n, y_0)]$, there follows by Lemma 3 in 4.3 the relation $d_e(y_n, y_0) = 0$. Thus d and d_e induce the same limit concept in E_2, and hence (see §2) it follows that $d \sim d_e$.

Case 2. $\lambda > 0$. Then we assert that $d \sim d_i$. Indeed, consider in E_2 any sequence x_n, $n = 0, 1, 2, \cdots$, such that $d(x_n, x_0) \to 0$ as $n \to \infty$. For n large enough, one has then

$$f[d_e(x_n, x_0)] = d(x_n, x_0) < \lambda.$$

Since $f(t) \geqq \lambda > 0$ for $t > 0$, it follows that $d_e(x_n, x_0) = 0$ and hence $x_n = x_0$ for n sufficiently large. Thus $d_i(x_n, x_0) = 0$ for n sufficiently large, and hence $d_i(x_n, x_0) \to 0$. Conversely, let y_n, $n = 0, 1, 2, \cdots$, be any sequence of points in E_2 such that $d_i(y_n, y_0) \to 0$. Then clearly $y_n = y_0$ and hence $d(y_n, y_0) = 0$ for n sufficiently large, and hence $d(y_n, y_0) \to 0$. Thus d and d_i induce the same limit concept, and hence (see Sec. 2) we have $d \sim d_i$.

4.5. In view of a remark in Section 2, we can introduce in E_2 the bounded distance

$$d_b(x, y) = \frac{d_e(x, y)}{1 + d_e(x, y)},$$

which satisfies the relations $0 \leqq d_b(x, y) < 1$, $d_b \sim d_e$. Accordingly, the theorem in 4.4 may be rephrased as follows: *if d is an invariant distance in E_2, then either $d \sim d_b$ or else $d \sim d_i$.*

5. Invariant distance-functions in G.

5.1. We now return to G, the set of all oriented lines in Euclidean three-space E_3. Our objective is to turn G into a metric space (G, δ) by selecting a suitable distance-function $\delta(g_1, g_2)$ of pairs of elements of G (see Sec. 2). If g_1, g_2 are elements of G, then their Euclidean distance $d_e(g_1, g_2)$ is defined by the formula

$$d_e(g_1, g_2) = \min d_e(x_1, x_2), \qquad x_1 \in [g_1], \qquad x_2 \in [g_2],$$

where $d_e(x_1, x_2)$ is the Euclidean distance of the points x_1, x_2. For each $g \in G$, let us denote by $u(g)$ the unit vector which is parallel to g and points in the direction of the oriented line g. It may seem natural to define then

$$e(g_1, g_2) = |u(g_2) - u(g_1)| + d_e(g_1, g_2)$$

as the distance of the oriented lines g_1, g_2, where $|u(g_2) - u(g_1)|$ is the length of the vector $u(g_2) - u(g_1)$. However, $e(g_1, g_2)$ is *not* a distance-function. Indeed, let g_1, g_3 be a pair of syntactic oriented lines such that $d_e(g_1, g_3) = 3/\sqrt{2}$, and let g_2 be an oriented line whose carrier $[g_2]$ intersects the carriers $[g_1]$, $[g_3]$ of g_1, g_3 at right angles. Then

$$e(g_1, g_2) = \sqrt{2}, \qquad e(g_2, g_3) = \sqrt{2}, \qquad e(g_1, g_3) = 3\sqrt{2},$$

and thus the triangle inequality $e(g_1, g_3) < e(g_1, g_2) + e(g_2, g_3)$ fails to hold.

5.2. It is now apparent that the selection of a suitable distance-function in G will require somewhat sophisticated thinking. To facilitate the discussion, we shall use vector notation in E_3, and the same letter x will denote the point of E_3 with coordinates x^1, x^2, x^3 and the vector with components x^1, x^2, x^3.

We shall use ρ as a generic notation for a rigid transformation in E_3, that is, a mapping $\rho: E_3 \to E_3$ such that $d_e[\rho(x), \rho(y)] = d_e(x, y)$ for every pair of points x, y in E_3.

If g is an oriented line and ρ is a rigid transformation in E_3, then the image $\rho(g)$ of g under ρ is defined as follows. Given an ordered pair (x, y) of distinct points in E_3, let us denote by $\gamma(x, y)$ the oriented line g which is (uniquely) determined by the following conditions: (i) the carrier $[g]$ of g contains the points x, y, and (ii) the direction from x to y on $[g]$ coincides with the orientation of g. Clearly, every oriented line can be represented in the form $g = \gamma(x, y)$ in infinitely many ways. Given now an oriented line $g = \gamma(x, y)$, we define $\rho(g) = \gamma[\rho(x), \rho(y)]$. It is immediate that $\rho(g)$ is independent of the particular choice of the representation $g = \gamma(x, y)$.

A distance-function $\delta(g_1, g_2)$ in G is *termed* invariant if

$$\delta[\rho(g_1), \rho(g_2)] = \delta(g_1, g_2)$$

for every choice of the oriented lines g_1, g_2 and of the rigid transformation ρ.

Our first objective is to obtain a view of the available supply of invariant distance-functions δ in G.

5.3. A first example of an invariant distance-function δ in G is the *isolating distance*

$$\delta_i(g_1, g_2) = \begin{cases} 0 \text{ if } g_1 = g_2, \\ 1 \text{ otherwise.} \end{cases}$$

A second example is the *semi-isolating distance*

$$\delta_{si}(g_1, g_2) = \begin{cases} \dfrac{d_e(g_1, g_2)}{1 + d_e(g_1, g_2)} \text{ if } g_1, g_2 \text{ are syntactic,} \\ 1 \text{ otherwise.} \end{cases}$$

The verification of the fact that δ_i, δ_{si} are actually invariant distance-functions in G is left to the reader. Let us merely note that if we had used $d_e(g_1, g_2)$ instead of $d_e(g_1, g_2)/[1 + d_e(g_1, g_2)]$ in defining δ_{si}, the resulting function would fail to satisfy the triangle inequality.

The determination of all invariant distance-functions δ in G is an open problem as yet. However, we are actually interested in the topologies $\Omega(\delta)$ induced in G by such invariant distance-functions, and the determination of all these topologies is a problem that can be readily solved. We shall prove, indeed, the following statement.

THEOREM. *If δ is any invariant distance-function in G, then either $\delta \sim \delta_i$ or else $\delta \sim \delta_{si}$.*

If we term the topology $\Omega(\delta_i)$ induced by δ_i the discrete topology, and if similarly we term the topology $\Omega(\delta_{si})$ induced by δ_{si} the semi-discrete topology in G, then the preceding theorem may be rephrased as follows: all the invariant distance-functions δ in G induce only two different topologies, namely, the discrete topology and the semi-discrete topology.

The proof of this theorem is based on a series of elementary lemmas which we shall consider presently.

5.4. Let δ be an invariant distance-function in G. Then δ gives rise to certain auxiliary quantities defined as follows. Let g_0', g_0'' be a pair of opposite oriented lines (see Sec. 1). We put $\Lambda_1 = \delta(g_0', g_0'')$. Then, since $g_0' \neq g_0''$, we have $0 < \Lambda_1 < \infty$. Furthermore, if g', g'' is any pair of opposite oriented lines, then $\delta(g', g'') = \Lambda_1$. Indeed, there exists a rigid transformation ρ in E_3 such that $\rho(g_0') = g'$, $\rho(g_0'') = g''$, and since δ is invariant it follows that

$$\delta(g', g'') = \delta[\rho(g_0'), \rho(g_0'')] = \delta(g_0', g_0'') = \Lambda_1.$$

Next, we consider a pair of oriented lines g', g'' whose carriers $[g']$, $[g'']$ intersect and are perpendicular to each other, and we put $\Lambda_2 = \delta(g', g'')$. Then we define, for $0 \leqq t < \infty$, two functions $f(t)$, $f^*(t)$ as follows. We select a pair of syntactic oriented lines g', g'' such that $d_s(g', g'') = t$, and we define $f(t) = \delta(g', g'')$. Next we select a pair of antitactic oriented lines g', g'' such that $d_s(g', g'') = t$ and we define $f^*(t) = \delta(g', g'')$. As in the case of Λ_1, we conclude readily from the invariance of δ that Λ_2, $f(t)$, $f^*(t)$ are independent of the particular choice of the pair g', g'', as specified in each of these cases. It is also clear that $0 < \Lambda_2 < \infty$, $0 \leqq f(t) < \infty$, $0 < f^*(t) < \infty$, $f^*(0) = \Lambda_1$. As regards $f(t)$, it is important to note that $f(t)$ is a scale-function for the Euclidean plane E_2. Indeed, let us think of E_2 as a horizontal plane in E_3, and for each point $x \in E_2$ let us denote by $g(x)$ the oriented line such that (i) the carrier of $g(x)$ contains x and is perpendicular to E_2, and (ii) $g(x)$ is oriented upward. Then

$$d(x_1, x_2) = \delta[g(x_1), g(x_2)]$$

is clearly a distance-function in E_2, and since $g(x_1)$ and $g(x_2)$ are syntactic and $d_s[g(x_1), g(x_2)] = d_s(x_1, x_2)$, it follows that

$$\delta[g(x_1), g(x_2)] = f[d_s(x_1, x_2)].$$

Thus $f[d_s(x_1, x_2)] = d(x_1, x_2)$ is a distance-function in E_2, and hence $f(t)$ is a scale-function for E_2. From Section 4, Lemmas 1 through 4, we obtain therefore the following information concerning $f(t)$.

LEMMA 1. *The function $f(t)$ has the following properties:*
(i) $0 \leqq f(t) < \infty$ *for* $0 \leqq t < \infty$.
(ii) $f(0) = 0$, $f(t) > 0$ *for* $t > 0$.

(iii) *If* (t_1, t_2, t_3) *is a Euclidean triple, then* $[f(t_1), f(t_2), f(t_3)]$ *is also a Euclidean triple.*

(iv) *If* t_n, $n=1, 2, \cdots$, *is any sequence such that* $0 \leqq t_n < \infty$ *and* $f(t_n) \to 0$, *then* $t_n \to 0$.

(v) *If* $0 \leqq t_1 \leqq t_2$, *then* $f(t_1) \leqq 2f(t_2)$.

(vi) *On setting*

$$\lambda = \underset{0<t<\infty}{\text{g.l.b.}} f(t), \qquad l = \liminf_{t \to +0} f(t), \qquad L = \limsup_{t \to +0} f(t),$$

we have the inequalities $\lambda \leqq l \leqq L \leqq 2\lambda$.

We proceed to state several further lemmas needed in the sequel.

LEMMA 2. $f^*(t) \geqq \Lambda_1/3$ *for* $0 \leqq t < \infty$.

LEMMA 3. *If* g', g'' *are a pair of oriented lines such that* $[g'] \cap [g''] \neq \varnothing$ *and* $[g'] \neq [g'']$, *then* $\delta(g', g'') \geqq f(1)/2$.

LEMMA 4. *If* g', g'' *are non-parallel oriented lines, then* $\delta(g', g'') \geqq f(1)/4$.

LEMMA 5. *If* g', g'' *are any two oriented lines, then* $\delta(g', g'') \leqq 2\Lambda_2$.

LEMMA 6. *Denote by* η *the smallest one of the quantities* $f(1)/4$, $\Lambda_1/3$, λ. *Then* $\delta(g', g'') > \eta$ *for every pair of oriented lines* g', g'' *such that* $g' \neq g''$.

LEMMA 7. *If* $\lambda = 0$, *then* $f(t)$ *is continuous at* $t=0$.

In the way of illustration, we indicate the proofs of the Lemmas 2, 3, 7, and we leave the analogous simple proofs of the Lemmas 4, 5, 6 to the reader. In proving Lemma 2, we note that $f^*(0) = \Lambda_1$, and hence it is sufficient to consider the case $t > 0$. Take a horizontal plane E_2 in E_3, and let P_1, P_2, P_3 be the vertices of an equilateral triangle of side-length t in E_2. Let g_1, g_1' be the oriented lines passing through P_1, P_2 respectively which are perpendicular to E_2 and are oriented upward, and let g_2, g_3 be the oriented lines passing through P_2, P_3 respectively which are perpendicular to E_2 and are oriented downward. Then

$$\Lambda_1 = \delta(g_1, g_1'), \qquad f^*(t) = \delta(g_2, g_1) = \delta(g_1, g_3) = \delta(g_3, g_1').$$

Thus the triangle inequality yields

$$\Lambda_1 = \delta(g_2, g_1') \leqq \delta(g_2, g_1) + \delta(g_1, g_3) + \delta(g_3, g_1') = 3f^*(t),$$

and Lemma 2 is proved (this concise argument was found by students of the writer in an introductory course in topology). To prove Lemma 3, denote by P the intersection of $[g']$ and $[g'']$, and take an oriented line g in the plane determined by $[g']$ and $[g'']$ such that g is syntactic to g' and $d_s(g', g) = 1$. Let Q be the intersection of $[g]$ and $[g'']$, and let ρ be the translation in E_3 which carries P into Q. Then $\rho(g') = g$, $\rho(g'') = g''$, and hence

$$\delta(g, g'') = \delta[\rho(g'), \rho(g'')] = \delta(g', g''),$$

since δ is invariant. It follows that

$$f(1) = \delta(g', g) \leqq \delta(g', g'') + \delta(g'', g) = 2\delta(g', g''),$$

and Lemma 3 is proved. As regards Lemma 7, the assumption $\lambda = 0$ implies (by part (vi) of Lemma 1) that $l = L = 0$, and since $f(0) = 0$, it follows that $f(t)$ is continuous at $t = 0$.

5.5. We are now ready to prove the theorem stated in 5.3. Consider an invariant distance δ in G.

Case 1. $\lambda > 0$. Then $\delta \sim \delta_i$. Indeed, consider any sequence g_n, $n = 0, 1, 2, \cdots$, of oriented lines such that $\delta(g_0, g_n) \to 0$. Since now the η in Lemma 6 is positive, we shall have $\delta(g_0, g_n) < \eta$ for n exceeding a properly chosen positive integer N. By Lemma 6 it follows that $g_n = g_0$ and hence $\delta_i(g_0, g_n) = 0$ for $n > N$. Thus $\delta_i(g_0, g_n) \to 0$. Conversely, consider any sequence g_n, $n = 0, 1, 2, \cdots$, of oriented lines such that $\delta_i(g_0, g_n) \to 0$. Then $\delta_i(g_0, g_n) = 0$ for n exceeding a properly chosen positive integer N, and hence $g_0 = g_n$ for $n > N$. Thus $\delta(g_0, g_n) = 0$ for $n > N$, and hence $\delta(g_0, g_n) \to 0$. It is thus established that $\Lambda(\delta) = \Lambda(\delta_i)$, and hence (see Sec. 2) we have shown that $\delta \sim \delta_i$.

Case 2. $\lambda = 0$. Then we assert that $\delta \sim \delta_{si}$. Indeed, consider any sequence g_n, $n = 0, 1, 2, \cdots$, of oriented lines. Assume first that $\delta_{si}(g_0, g_n) \to 0$. Then $\delta_{si}(g_0, g_n) < 1$ for n exceeding a properly chosen positive integer N, and hence g_n and g_0 are syntactic and

$$\delta_{si}(g_0, g_n) = \frac{d_s(g_0, g_n)}{1 + d_s(g_0, g_n)} \quad \text{for } n > N.$$

Since $\delta_{si}(g_0, g_n) \to 0$, it follows that

$$d_s(g_0, g_n) \to 0. \tag{1}$$

Since g_0 and g_n are syntactic, we have also

$$\delta(g_0, g_n) = f[d_s(g_0, g_n)] \quad \text{for } n > N. \tag{2}$$

Since now $\lambda = 0$, by Lemma 7 it follows from (1) and (2) that $\delta(g_0, g_n) \to 0$.

Assume, conversely, that $\delta(g_0, g_n) \to 0$. Then $\delta(g_0, g_n)$ will be less than the smaller one of $f(1)/4$, $\Lambda_1/3$ for n exceeding a properly chosen positive integer N. From the Lemmas 4 and 2 we conclude that g_0 and g_n are syntactic for $n > N$, and hence

$$\delta(g_0, g_n) = f[d_s(g_0, g_n)] \quad \text{for } n > N, \tag{3}$$

$$\delta_{si}(g_0, g_n) = \frac{d_s(g_0, g_n)}{1 + d_s(g_0, g_n)} \quad \text{for } n > N. \tag{4}$$

Since $\delta(g_0, g_n) \to 0$, it follows from (3) and Lemma 1 that $d_s(g_0, g_n) \to 0$. Hence $\delta_{si}(g_0, g_n) \to 0$ in view of (4). It is thus established that $\Lambda(\delta) = \Lambda(\delta_{si})$, and hence $\delta \sim \delta_{si}$.

5.6. In view of the theorem just proved, $\Omega(\delta_i)$ and $\Omega(\delta_{si})$ are the only topologies in G that are induced by invariant distance-functions δ in G. Unfortunately, neither one of these topologies is satisfactory in relation to most applications. To illustrate this point, consider an oriented circle C in a plane E_2, and denote by $g(x)$ the correspondingly oriented tangent of C at the point x of C. As a matter of both intuition and practical necessity, the oriented line $g(x)$ should vary continuously with the point x of C. Consider now a sequence x_n, $n=0, 1, 2, \cdots$, of distinct points of C such that $d_e(x_0, x_n)\to 0$. Then g_0 and g_n are distinct and not parallel, and hence $\delta_i[g(x_0), g(x_n)]=\delta_{si}[g(x_0), g(x_n)]=1$, $n=1, 2, \cdots$. In view of the theorem in 5.3 it follows that the relation $g(x_n)\to g(x_0)$ fails to hold in every topology induced by an invariant distance-function in G. Accordingly, a different approach is needed to obtain a useful topology in G.

6. G as a limit space.

6.1. Let us recall (see 5.2) that every oriented line g in E_3 can be represented, in infinitely many ways, in the form $g=\gamma(x, y)$, where (x, y) is an ordered pair of distinct points on $[g]$ such that the direction from x toward y on $[g]$ agrees with the orientation of g. Given then a sequence g_n, $n=0, 1, 2, \cdots$, of oriented lines in G, we shall say that g_n converges to g_0, in symbols $g_n\to g_0$, if and only if there exist representations $g_n=\gamma(x_n, y_n)$, $n=0, 1, 2, \cdots$, such that $d_e(x_0, x_n)\to 0$, $d_e(y_0, y_n)\to 0$. This limit concept in G, to be denoted by Λ, is motivated by the requirement that the oriented line $g=\gamma(x, y)$ should depend continuously upon the points x, y. In fact, this limit concept Λ is the one used (explicitly or implicitly) in all major applications at present. This limit concept Λ gives rise to the following observations.

I. Λ is invariant under rigid transformations in E_3. That is, if $g_n\to g_0$, then (see 5.2) we have also $\rho(g_n)\to\rho(g_0)$ for every rigid transformation ρ in E_3. The proof is obvious.

II. The limit space (G, Λ) is metrizable (that is, there exists a distance-function $d(g_1, g_2)$ in G such that $\Lambda=\Lambda(d)$). Indeed, let o be the origin in E_3, and for each oriented line g in E_3 denote by $\phi(g)$ the point on g nearest to o. Let again $u(g)$ be the unit vector which agrees in direction and orientation with g. On setting

$$d(g_1, g_2) = |u(g_2) - u(g_1)| + |\phi(g_2) - \phi(g_1)|,$$

it is immediate that we obtain a distance-function d in G such that $\Lambda=\Lambda(d)$.

III. There exists no invariant distance-function δ in G such that $\Lambda=\Lambda(\delta)$ (thus, even though Λ is both invariant and metrizable, it cannot be metrized in terms of an invariant distance-function). Indeed, returning to the circle C considered in 5.6, we verify readily that $g(x_n)\to g(x_0)$ under the present limit concept Λ, while (as noted in 5.6), the relation $g(x_n)\to g(x_0)$ fails to hold for every limit concept $\Lambda(\delta)$ where δ is an invariant distance-function in G.

6.2. It is immediate that the distance-function d introduced in 6.1, II, is not invariant under rigid transformations in E_3 (and in fact it cannot be invariant in view of 6.1, III). Accordingly, there arise the following two questions concerning the topology $\Omega(d)$ induced by d.

(i) Is $\Omega(d)$ invariant under rigid transformations in E_3 (that is, if $O \in \Omega(d)$, is it true that $\rho(O) \in \Omega(d)$ for every rigid transformation ρ in E_3)?

(ii) Is the topology $\Omega(d)$ uniquely determined by the limit concept Λ that we adopted for G?

The affirmative answers follow readily from the general comments made in Section 2. Indeed, if d', d'' are any two distance-functions on G such that $\Lambda(d') = \Lambda = \Lambda(d'')$, then (see Sec. 2) it follows that $\Omega(d') = \Omega(d'')$, and thus question (ii) is answered in the affirmative. Returning to d, we note that a set $E \subset G$ is closed in the topology $\Omega(d)$ if and only if the relations $g_n \in E$, $n = 1, 2, \cdots$, and $g_n \to g_0$ imply that $g_0 \in E$. Since the limit concept $\Lambda = \Lambda(d)$ is invariant under rigid transformations in E_3, it follows that if E is closed in the topology $\Omega(d)$ then $\rho(E)$ is also closed, for every rigid transformation ρ in E_3. The corresponding statement for open sets follows now directly, since the complement of an open set is a closed set.

6.3. In summary, the fruitful and generally used limit concept Λ described in 6.1 is invariant under rigid transformations. On defining closed sets in G as those sets which contain their limit elements, and on introducing open sets in G as sets whose complements are closed, we obtain a corresponding topology Ω in G which is both invariant and metrizable, but is not metrizable in terms of an invariant distance-function. All the invariant distance-functions in G induce only two distinct topologies. These two topologies, which we termed the discrete and the semidiscrete topology respectively, are useless as far as present major applications are concerned.

A COVERING THEOREM FOR UNIVALENT FUNCTIONS

W. T. SCOTT, Northwestern University

Let U denote the class of normalized univalent functions $f(z)$ of the form

$$f(z) = z + a_2z^2 + \cdots + a_nz^n + \cdots, \qquad |z| < 1.$$

An elementary yet interesting and important property shared by functions of U is stated in the Bieberbach coefficient theorem [1]:

If $f(z)$ belongs to U then

$$|a_2| \leqq 2;$$

moreover, $|a_2| = 2$ for a function $f(z)$ of U if and only if

$$f(z) = k_\alpha(z) = \frac{z}{(1 - e^{i\alpha}z)^2}, \qquad |z| < 1, \ \alpha \text{ real}.$$

An immediate consequence of the above result is the well known Koebe-Bieberbach covering theorem [1, 2] (see, for example, [3, pp. 209 ff.]):

If $f(z)$ belongs to U and omits the value γ, then

$$|\gamma| \geqq 1/4;$$

moreover, $\gamma = -e^{-i\alpha}/4$, α real, is an omitted value for $f(z)$ of U if and only if $f(z) = k_\alpha(z)$.

This theorem shows that the map of $|z| < 1$ by a function of U always contains the open disc $|w| < 1/4$, and that there is no larger disc with center at $w = 0$ which is contained in every such map. For contrast with the result to be presented below, we emphasize the fact that any boundary point of the disc $|w| < 1/4$ corresponds to an omitted value for some function of U.

Let U_+ denote the subclass of U composed of the identity function, z, and those functions of U for which the first non-zero power series coefficient, a_m, $m \geqq 2$, satisfies $a_m > 0$. Since for real θ, $e^{i\theta}f(e^{-i\theta}z)$ belongs to U whenever $f(z)$ belongs to U, it follows that to each function of U there corresponds by such a transformation a uniquely determined function of U_+.

For $f(z)$ in U_+ and for fixed ϕ let $\rho_f(\phi)$ denote the distance, perhaps infinite, from $w = 0$, along the ray $\arg w = \phi$, to the nearest boundary point of the map of $|z| < 1$ by $w = f(z)$. Since the function $f(z) = z$ is in U_+,

$$\rho(\phi) = \operatorname*{glb}_{f \in U_+} \rho_f(\phi)$$

is finite and, because of the Koebe-Bieberbach theorem, satisfies

$$1/4 \leqq \rho(\phi) \leqq 1.$$

From the definition of $\rho(\phi)$ it is evident that if $\gamma=\rho e^{i\phi}$ is an omitted value for a function of U_+, then $\rho\geqq\rho(\phi)$. It is also evident that $\rho(-\phi)=\rho(\phi)$, since the function

$$\tilde{f}(z) = z + \bar{a}_2 z^2 + \cdots + \bar{a}_n z^n + \cdots, \qquad |z| < 1,$$

belongs to U_+ whenever $f(z)$ belongs to U_+, and the values taken on by $\tilde{f}(z)$ in $|z|<1$ are the conjugates of values taken on by $f(z)$ in $|z|<1$.

THEOREM. *If $f(z)$ belongs to U_+ and for $|z|<1$ omits the value $\gamma=\rho e^{i\phi}$ then $\rho\geqq\rho(\phi)$, and*

$$\begin{aligned} \rho(\phi) &= 1/2, \qquad 0 \leqq |\phi| \leqq \pi/2, \\ \tfrac{1}{2}|\sin\phi| < \rho(\phi) &\leqq \tfrac{1}{2}, \qquad \pi/2 < |\phi| \leqq 3\pi/4, \\ \frac{1}{4|\cos\phi|} < \rho(\phi) &\leqq \tfrac{1}{2}, \qquad 3\pi/4 \leqq |\phi| < \pi, \\ \rho(+\pi) &= 1/4. \end{aligned}$$

Moreover, for $0\leqq|\phi|<\pi/2$, $\rho(\phi)e^{i\phi}$ is not an omitted value for any function of U_+, but for $\pi/2\leqq|\phi|\leqq\pi$, each value $\rho(\phi)e^{i\phi}$ is an omitted value for some function of U_+.

We start the proof by deducing a necessary condition for a function $f(z)$ of U_+ to omit the value $\gamma=\rho e^{i\phi}$. The function

$$g(z) = \frac{f(z)}{1-f(z)/\gamma} = z + \left(a_2 + \frac{1}{\gamma}\right)z^2 + \cdots$$

belongs to U and hence $|a_2+1/\gamma|\leqq 2$, or, in an equivalent statement,

$$(4-a_2^2)\rho^2 - 2a_2\rho\cos\phi - 1 \geqq 0,$$

and equality can hold here only if $g(z)=k_\alpha(z)$ for suitable real α.

Geometric motivation for the continuation of the proof may be had by using equality in the above necessary condition and regarding a_2, $0\leqq a_2\leqq 2$, as a parameter. It is readily found that the interiors of the resulting circles have in common a region on whose boundary the value of ρ is precisely that given as a lower bound for $\rho(\phi)$ in the theorem.

We return to the proof of the theorem and suppose that $\rho\leqq 1/2$, $0\leqq|\phi|\leqq\pi/2$. Then $(4-a_2^2)\rho^2-2a_2\rho\cos\phi-1=-(1-4\rho^2)-a_2\rho(a_2\rho+2\cos\phi)\leqq 0$, and in view of the stated necessary condition it follows that $\rho=1/2$, $a_2=0$, $g(z)=k_\alpha(z)$ and

$$f(z) = \frac{z}{1-2(e^{i\alpha}-e^{-i\phi})z+e^{2i\alpha}z^2}.$$

We observe that $\alpha=-\phi$ since $a_2=0$, and conclude thereafter that $\phi=\pm\pi/2$ since $a_3=-e^{-2i\phi}$ must be positive. Thus we have shown that $\rho(\phi)\geqq 1/2$ for $0\leqq|\phi|\leqq\pi/2$ and that no function of U_+ omits the value $\frac{1}{2}e^{i\phi}$ for $0\leqq|\phi|<\pi/2$.

The function $f(z)$ obtained above is $z/(1-z^2)$, which omits the values $\pm i/2$, and, in particular, $\rho(\pm\pi/2)=1/2$.

Suppose next that $\rho \leqq |\sin\phi|/2$, $\pi/2<|\phi| \leqq 3\pi/4$. Then

$$(4-a_2^2)\rho^2 - 2a_2\rho\cos\phi - 1 = 4\rho^2 - \sin^2\phi - (a_2\rho+\cos\phi)^2 \leqq 0,$$

from which it follows that $\rho=|\sin\phi|/2$, $a_2\rho+\cos\phi=0$, and hence that $a_2=-2\cos\phi/|\sin\phi|$. In addition, $g(z)=k_\alpha(z)$, or

$$f(z) = \frac{z}{1-2(e^{i\alpha}-e^{-i\phi}/|\sin\phi|)z+e^{2i\alpha}z^2},$$

and

$$a_2 = 2\left[\cos\alpha - \frac{\cos\phi}{|\sin\phi|} + i\left(\sin\alpha + \frac{\sin\phi}{\sin|\phi|}\right)\right].$$

Comparison of the values of a_2 gives $\phi=-\alpha=\pm\pi/2$, which is not in the permissible range of ϕ. Thus no function of U_+ omits $\rho e^{i\phi}$ for $\rho \leqq |\sin\phi|/2$, $\pi/2<|\phi| \leqq 3\pi/4$, and $\rho(\phi) \geqq |\sin\phi|/2$ in this range.

Now suppose that $\rho \leqq 1/4|\cos\phi|$, $3\pi/4 \leqq |\phi| \leqq \pi$. Since $\cos^2\phi \geqq 1/2$,

$$(4-a_2^2)\rho^2 - 2a_2\rho\cos\phi - 1 \leqq -\frac{(2-a_2)(8\cos^2\phi-2-a_2)}{16\cos^2\phi} \leqq 0,$$

and it follows from the necessary condition that $a_2=2$, $f(z)=z/(1-z)^2$, and $\gamma=-1/4$. Thus $\rho(\pm\pi)=1/4$, $\rho(\phi) \geqq 1/4|\cos\phi|$ for $3\pi/4 \leqq |\phi|<\pi$, and no function of U_+ omits $\rho e^{i\phi}$ for $\rho \leqq 1/4|\cos\phi|$, $3\pi/4 \leqq |\phi|<\pi$.

Next we show that for any β, $0<|\beta|<\pi/2$, there exists a function of U_+ which omits the value $te^{i\beta}$, where t, which is necessarily greater than $1/2$, is arbitrarily close to $1/2$. This will complete the proof that $\rho(\phi)=1/2$, $0 \leqq |\phi| \leqq \pi/2$.

The function $f(z)=z/(1+e^{-2i\beta}z^2)$ belongs to U and omits the values $e^{i\beta}/2$, $e^{i(\beta+\pi)}/2$. It is readily seen that for $|\lambda|<1$,

$$h(z,\lambda) = \frac{1}{(1-\lambda\bar{\lambda})f'(\lambda)}\left[f\left(\frac{z+\lambda}{\bar{\lambda}z+1}\right) - f(\lambda)\right]$$

belongs to U, and by computation it is found that

$$h(z,\lambda) = \frac{z + \dfrac{\bar{\lambda}-e^{-2i\beta}\lambda}{1-e^{-2i\beta}\lambda^2}}{1+2\dfrac{\bar{\lambda}+e^{-2i\beta}\lambda}{1+e^{-2i\beta}\lambda^2}z + \dfrac{\bar{\lambda}^2+e^{-2i\beta}}{1+e^{-2i\beta}\lambda^2}z^2}$$
$$= z + A(\lambda)e^{i\alpha(\lambda)}z^2 + \cdots,$$

where

$$A(\lambda)e^{i\alpha(\lambda)} = e^{-i\beta}\left[\frac{e^{i\beta}\bar{\lambda} - e^{-i\beta}\lambda}{1 - e^{-2i\beta}\lambda^2} - 2\frac{e^{i\beta}\bar{\lambda} + e^{-i\beta}\lambda}{1 + e^{-2i\beta}\lambda^2}\right].$$

If λ is chosen so that $A(\lambda)>0$, then the function

$$f(z, \lambda) = e^{i\alpha(\lambda)}h[e^{-i\alpha(\lambda)}z, \lambda] = z + A(\lambda)z^2 + \cdots$$

belongs to U_+. Another computation shows that the omitted values of $f(z, \lambda)$ corresponding to the omitted values $\pm e^{i\beta}/2$ of $f(z)$ are

$$\frac{e^{i\beta}}{2}\cdot e^{i\alpha(\lambda)}\frac{(1 + e^{-2i\beta}\lambda^2)(1 - e^{-i\beta}\lambda)}{(1 - \lambda\bar{\lambda})(1 + e^{-i\beta}\lambda)}, \qquad -\frac{e^{i\beta}}{2}\cdot e^{i\alpha(\lambda)}\frac{(1 + e^{-2i\beta}\lambda^2)(1 + e^{-i\beta}\lambda)}{(1 - \lambda\bar{\lambda})(1 - e^{-i\beta}\lambda)}.$$

For $\lambda = ce^{i\gamma}$ power series expansion gives

$$X = A(\lambda)e^{i\alpha(\lambda)}\frac{(1 + e^{-2i\beta}\lambda^2)(1 - e^{-i\beta}\lambda)}{1 + e^{-i\beta}\lambda} = c[-e^{-i\gamma} - 3e^{i(\gamma-2\beta)} + O(c)],$$

and it is readily found that a value γ_0 of γ for which

$$\operatorname{Im}\left[-e^{-i\gamma} - 3e^{i(\gamma-2\beta)}\right] = \sin\gamma - 3\sin(\gamma - 2\beta) = 0$$

is given by

$$\sin\gamma_0 = -\frac{3\sin 2\beta}{\sqrt{10 - 6\cos 2\beta}}, \qquad \cos\gamma_0 = \frac{1 - 3\cos 2\beta}{\sqrt{10 - 6\cos 2\beta}}.$$

For $\gamma = \gamma_0$, $\operatorname{Re}\left[-e^{-i\gamma} - 3e^{i(\gamma-2\beta)}\right] > 0$ and, because of the continuity of the function X, for $c>0$, sufficiently small, there is a closed interval about γ_0 in which $\operatorname{Re} X > 0$. For γ in this closed interval $\operatorname{Im}\left[-e^{-i\gamma} - 3e^{i(\gamma-2\beta)}\right]$ changes sign and hence, for $c>0$, sufficiently small, there is a value γ_1 for which $\operatorname{Im} X = 0$. Thus for each $c>0$, sufficiently small, there is a γ_1 for which $\lambda = ce^{i\gamma_1}$ gives $A(\lambda)>0$ and

$$\arg e^{i\alpha(\lambda)}\frac{(1 + e^{-2i\beta}\lambda^2)(1 - e^{-i\beta}\lambda)}{1 + e^{-i\beta}\lambda} = 0.$$

That is, for $\lambda = ce^{i\gamma_1}$

$$\frac{e^{i\beta}}{2}\left|\frac{(1 + e^{-2i\beta}\lambda^2)(1 - e^{-i\beta}\lambda)}{(1 - \lambda\bar{\lambda})(1 + e^{-i\beta}\lambda)}\right| = te^{i\beta}$$

is an omitted value for $f(z, \lambda)$. For any β, $0 \leqq |\beta| \leqq \pi$ it is clear that $t \to 1/2$ as $c \to 0$ and it follows that $\rho(\phi) \leqq 1/2$, $0 \leqq |\phi| \leqq \pi$.

There remains to be proved the statement that each value $\rho(\phi)e^{i\phi}$, $\pi/2 \leqq |\phi| \leqq \pi$, is an omitted value for some function of U_+. It is sufficient to consider the interval $\pi/2 < |\phi| < \pi$ since $\rho(\pm\pi/2) = 1/2$, $\rho(\pm\pi) = 1/4$, and the values $\pm i/2$, $-1/4$, are omitted by the functions $z/(1-z^2)$, $z/(1-z)^2$, respectively. Also, we

need consider only the case $1/4<\rho(\phi)<1/2$, $\pi/2<|\phi|<\pi$ since the function $f(z, \lambda)$ obtained by putting $\beta=-\pi/2$, $\lambda=c$, $0<c<1$, is

$$\left(z+\frac{2c}{1+c^2}z^2\right)\Big/(1-z^2),$$

and this function of U_+ omits the values

$$\frac{1}{2}\left(-\frac{2c}{1+c^2}\pm i\frac{1-c^2}{1+c^2}\right)$$

whose modulus is $1/2$.

From the definition of $\rho(\phi)$ we see that either there is a function of U_+ which omits $\rho(\phi)e^{i\phi}$, or else there is a sequence of functions $\{f_n(z)\}$ of U_+ for which $\rho_{f_n}(\phi)>\rho(\phi)$, $\rho_{f_n}(\phi)\to\rho(\phi)$. In the latter case, the sequence $\{f_n(z)\}$ has a limit function $F(z)$ in U since $U_+\subset U$ and U is a compact normal family; moreover there is a subsequence of $\{f_n(z)\}$ which converges uniformly to $F(z)$ in any closed subset of $|z|<1$. From the inequality $|a_2+1/\gamma|\leqq 2$ we get $a_2\geqq -2+1/|\gamma|$, and it follows that the coefficients a_2 for functions of the subsequence are ultimately bounded away from 0. This enables us to conclude that $F(z)$ belongs to U_+. The assumption that $F(z_0)=\rho(\phi)e^{i\phi}$ for some z_0 in $|z|<1$ together with the uniform convergence of the subsequence to $F(z)$ in a sufficiently small closed disc about z_0 leads to the conclusion that the functions of the subsequence do not omit values having $\rho(\phi)e^{i\phi}$ as a limit point, and it follows that $F(z)$ omits $\rho(\phi)e^{i\phi}$.

This completes the proof of the theorem since the possibility that $\rho(\phi)$ equals the stated lower bound for $\pi/2<|\phi|<\pi$ is now excluded.

It is worth noting that the region bounded by the curve $w=\rho(\phi)e^{i\phi}$ is not convex in the neighborhood of $w=-1/4$. This fact is an immediate consequence of the inequality $\rho(\phi)\cos\phi<-1/4$, $3\pi/4\leqq|\phi|<\pi$.

We remark in conclusion that the bound $\rho(\phi)\leqq 1/2$, $\pi/2<|\phi|<\pi$, can be improved by minimizing the modulus of an omitted value of $f(z, \lambda)$, subject to the condition that the argument of the omitted value shall be ϕ. This procedure has not been attempted since it appears to be very tedious and shows no promise of yielding the precise value of $\rho(\phi)$.

References

1. L. Bieberbach, Über die Koeffizienten derjenigen Potenzreihen, welche eine schlichte Abbildung des Einheitskreises vermitteln, K. Preuss. Akad. Wiss. Sitzungsberichte, 1916, pp. 940–955.

2. P. Koebe, Über die Uniformisierung belibiger analytischer Kurven, Nachr. Ges. Wiss. Göttingen, 1907, pp. 191–210.

3. Nehari, Conformal Mapping, McGraw-Hill, New York, 1952.

SOME CONVERGENCE PROBLEMS FOR CONTINUED FRACTIONS

H. S. WALL, The University of Texas

We discuss in this paper the problem of finding conditions for a continued fraction to converge in case it is given that certain subsequences of the sequence of approximants converge (Sec. 2 and Sec. 3). Use is made of formulas for cross-ratios of four approximants (Sec. 1). These formulas were used in [6] to obtain continued fractions f' whose sequences of approximants are rearrangements of the sequence of approximants of the continued fraction f. By applying known convergence criteria to f' we found new criteria for convergence of f. We are now able to find several interesting theorems in this way by means of convergence criteria discovered since [6] was written (Sec. 4 and Sec. 5).

1. Cross-ratios. Suppose a is the complex number sequence $\{a_p\}_{p=1}^{\infty}$, no term of which is 0, q a nonnegative integer, $a^q=\{a_{q+p}\}_{p=1}^{\infty}$ and $f_p^q=A_p^q/B_p^q$ the pth approximant of the continued fraction

$$f(a^q)=\frac{1}{1}+\frac{a_1^q}{1}+\frac{a_2^q}{1}+\cdots=\frac{1}{1}+\frac{a_{q+1}}{1}+\frac{a_{q+2}}{1}+\cdots,$$

so that

$$\begin{aligned} &A_0^q=0,\ A_1^q=1,\ A_{p+1}^q=A_p^q+a_{q+p}A_{p-1}^q,\\ &B_0^q=1,\ B_1^q=1,\ B_{p+1}^q=B_p^q+a_{q+p}B_{p-1}^q, \end{aligned} \qquad p=1,2,\cdots. \tag{1.1}$$

We omit the superscript q in case q is 0. An easy induction argument shows that $A_{p+1}^q=B_p^{q+1}$, $p, q=0, 1, \cdots$; and also, if $a_0=1$,

$$A_{p+k}B_p-A_pB_{p+k}=(-1)^p a_0a_1\cdots a_pA_k^p, \quad p, k=0, 1, \cdots. \tag{1.2}$$

If $k=1$, (1.2) shows that $A_p\neq 0$ or $B_p\neq 0$, so that f_p is a point in the extended complex plane. If p and q are integers, $0\leqq p<q$, $B_p\neq 0$ and $B_q\neq 0$ then, by (1.2),

$$f_q-f_p=\frac{(-1)^p a_0a_1\cdots a_pA_{q-p}^p}{B_pB_q}. \tag{1.3}$$

Therefore, if p, q, r and s are integers and $0\leqq p<q<r<s$,

$$[f_s, f_q, f_p, f_r]=\frac{(f_s-f_q)(f_p-f_r)}{(f_s-f_p)(f_q-f_r)}=\frac{A_{s-q}^qA_{r-p}^p}{A_{s-p}^pA_{r-q}^q}, \tag{1.4}$$

provided this cross-ratio exists. If $L=[f_p, f_q, f_r, f_s]$, then

$$\begin{aligned} &L=[f_q, f_p, f_s, f_r]=[f_r, f_s, f_p, f_q]=[f_s, f_r, f_q, f_p],\\ &\frac{1}{L}=[f_p, f_r, f_q, f_s] \quad \text{and} \quad 1-L=[f_s, f_p, f_q, f_r]. \end{aligned} \tag{1.5}$$

Then, by (1.4),

$$[f_p, f_q, f_r, f_s] = 1 - \frac{A^p_{s-p} A^q_{r-q}}{A^p_{r-p} A^q_{s-q}}. \tag{1.6}$$

If, in particular, $q=p+1$, $r=p+2$ and $s=p+3$, this reduces to:

$$a_{p+2} = -[f_p, f_{p+1}, f_{p+2}, f_{p+3}], \qquad p = 0, 1, \cdots. \tag{1.7}$$

2. Continued fractions for which $\{f_{2p-1}\}$ and $\{f_{2p}\}$ converge. The statement that the sequence $\{x_p\}_{p=1}^{\infty}$ is absolutely convergent means the series $x_1 + \sum_{p=1}^{\infty}(x_{p+1}-x_p)$ is absolutely convergent.

THEOREM 2.1. *Each of the following conditions is sufficient for convergence of the continued fraction $f(a)$:*

(1) *$\{f_{2p-1}\}$ and $\{f_{2p}\}$ converge and a has a bounded infinite subsequence* [7],

(2) *$\{f_{2p-1}\}$ and $\{f_{2p}\}$ converge, one of these sequences converges absolutely, and the series $\sum|a_p|^{-1}$ diverges, and*

(3) *$\{f_{2p-1}\}$ and $\{f_{2p}\}$ converge absolutely and the series $\sum|a_p|^{-1/2}$ diverges.*

Proof. This is an easy consequence of the formula (1.7), *i.e.*,

$$a_{p+2}\cdot(f_p - f_{p+2})(f_{p+1} - f_{p+3}) = -(f_p - f_{p+1})(f_{p+2} - f_{p+3}). \tag{2.1}$$

If $\{f_{2p-1}\}$ and $\{f_{2p}\}$ converge, a has a bounded infinite subsequence and $f(a)$ does not converge, there exists a positive number c and a positive integer N such that $|(f_p - f_{p+1})(f_{p+2} - f_{p+3})| > c$, if $p > N$, and an integer p greater than N such that $|a_{p+2}(f_p - f_{p+2})(f_{p+1} - f_{p+3})| < c$. This contradicts (2.1). If $\{f_{2p-1}\}$ and $\{f_{2p}\}$ converge, one absolutely, then $\sum|(f_p - f_{p+2})(f_{p+1} - f_{p+3})|$ converges and, by (2.1), $\sum|a_{p+2}|^{-1}|(f_p - f_{p+1})(f_{p+2} - f_{p+3})|$ converges. Since $\sum|a_{p+2}|^{-1}$ diverges, it follows that $f(a)$ converges. Finally, if $\{f_{2p-1}\}$ and $\{f_{2p}\}$ converge absolutely, we see, by (2.1), that $\sum|a_{p+2}|^{-1/2}|(f_p - f_{p+1})(f_{p+2} - f_{p+3})|^{1/2} \leqq \frac{1}{2}\sum\{|f_p - f_{p+2}| + |f_{p+1} - f_{p+3}|\}$ so that $f(a)$ converges in case $\sum|a_{p+2}|^{-1/2}$ diverges.

In [2], R. E. Lane and the writer showed that, if $\{f_{2p-1}\}$ and $\{f_{2p}\}$ converge absolutely, then $f(a)$ converges only in case the series $\sum|b_p|$ diverges, where

$$b_1 = 1, \qquad b_{p+1} = 1/a_p b_p, \qquad p = 1, 2, \cdots. \tag{2.2}$$

David F. Dawson recently* obtained the following more comprehensive result:

If $\{b_p\}_{p=1}^{\infty}$ is a complex number sequence, the continued fraction

$$\frac{1}{b_1} + \frac{1}{b_2} + \frac{1}{b_3} + \cdots \tag{2.3}$$

converges in case

* Unpublished.

(1) $\{f_{2p-1}\}$ *converges absolutely,* $\{f_{2p}\}$ *converges and* $\sum|b_{2p-1}|$ *diverges, or*
(2) $\{f_{2p-1}\}$ *converges,* $\{f_{2p}\}$ *converges absolutely and* $\sum|b_{2p}|$ *diverges.*

In [7] we gave an example of a divergent continued fraction $f(a)$ such that $\{f_{2p-1}\}$ and $\{f_{2p}\}$ converge and $\sum|a_p|^{-1/2}$ (and hence $\sum|b_p|$) diverges. For that example, $a_p=(-1)^p(p+1)^2$ so that $\sum|a_p|^{-1}$ converges. *There exists a divergent continued fraction* $f(a)$ *such that* $\{f_{2p-1}\}$ *and* $\{f_{2p}\}$ *converge and* $\sum|a_p|^{-1}$ *diverges.* In fact, if $f_0=0$,

$$f_{2p} = -1^{-1/2} + 2^{-1/2} - \cdots + (-1)^p p^{-1/2}, \qquad p = 1, 2, \cdots$$

and $f_{2p+1}=1+f_{2p}$, $p=0, 1, \cdots$, then $\{f_{2p-1}\}$ and $\{f_{2p}\}$ converge, $\{f_p\}$ diverges and, by (2.1), $a_{2p+2}=-(p+1)$, so that $\sum|a_p|^{-1}$ diverges. Note that the four sequences $\{f_{4p+1}\}$, $\{f_{4p+2}\}$, $\{f_{4p+3}\}$ and $\{f_{4p+4}\}$ are absolutely convergent.

3. The sets M_k. For each positive integer k, M_k denotes the set of all continued fractions $f(a)$ for which there exists a number k-tuple $L_1, \cdots, L_k$ such that $\lim_{p\to\infty} f_{kp+r}=L_r$, $r=1, \cdots, k$.

THEOREM 3.1. *If the continued fraction* $f(a)$ *belongs to* M_3 *and none of the sequences* $\{a_{3p}\}$, $\{a_{3p+1}\}$ *or* $\{a_{3p+2}\}$ *converges to the limit* -1, *then* $f(a)$ *converges.*

Proof. By (1.5) and (1.7),

$$(f_{p+1} - f_{p+3})(f_{p+2} - f_p) = \frac{1}{1 + a_{p+2}}(f_{p+1} - f_{p+2})(f_{p+3} - f_p).$$

If $f(a)$ belongs to M_3 and none of the sequences $\{a_{3p}\}$, $\{a_{3p+1}\}$ or $\{a_{3p+2}\}$ converges to the limit -1, it follows that each of the following statements is true: (1) $L_1=L_3$ or $L_2=L_3$, (2) $L_2=L_1$ or $L_3=L_1$ and (3) $L_3=L_2$ or $L_1=L_2$. Consequently, $L_1=L_2=L_3$ so that $f(a)$ is convergent.

The continued fraction $f(a)$ diverges if $a_p=-1$, $p=1, 2, 3, \cdots$. More generally, we have this theorem:

THEOREM 3.2. *The continued fraction* $f(a)$ *diverges in case any one of the following series converges:*

$$\left|\frac{1}{a_0a_1}(1+a_1)\right| + \left|\frac{a_0a_1}{a_2}(1+a_3)\right| + \left|\frac{a_2}{a_0a_1\cdot a_3a_4}(1+a_4)\right|$$

$$+ \left|\frac{a_0a_1\cdot a_3a_4}{a_2\cdot a_5}(1+a_6)\right| + \left|\frac{a_2\cdot a_5}{a_0a_1\cdot a_3a_4\cdot a_6a_7}(1+a_7)\right| \tag{3.1}$$

$$+ \left|\frac{a_0a_1\cdot a_3a_4\cdot a_6a_7}{a_2\cdot a_5\cdot a_8}(1+a_8)\right| + \cdots,$$

$$\left|\frac{1}{a_0}(1+a_1)\right|+\left|\frac{a_0}{a_1a_2}(1+a_2)\right|+\left|\frac{a_1a_2}{a_0\cdot a_3}(1+a_4)\right|+\left|\frac{a_0\cdot a_3}{a_1a_2\cdot a_4a_5}(1+a_5)\right| \tag{3.2}$$
$$+\left|\frac{a_1a_2\cdot a_4a_5}{a_0\cdot a_3\cdot a_6}(1+a_7)\right|\left|\frac{a_0\cdot a_3\cdot a_6}{a_1a_2\cdot a_4a_5\cdot a_7a_8}(1+a_8)\right|+\cdots,$$

or

$$\left|\frac{1}{a_1}(1+a_2)\right|+\left|\frac{a_1}{a_2a_3}(1+a_3)\right|\quad\left|\frac{a_2a_3}{a_1\cdot a_4}(1+a_5)\right|+\left|\frac{a_1\cdot a_4}{a_2a_3\cdot a_5a_6}(1+a_6)\right| \tag{3.3}$$
$$+\left|\frac{a_2a_3\cdot a_5a_6}{a_1\cdot a_4\cdot a_7}(1+a_8)\right|+\left|\frac{a_1\cdot a_4\cdot a_7}{a_2a_3\cdot a_5a_6\cdot a_8a_9}(1+a_9)\right|+\cdots.$$

Proof. The continued fractions

$$1-\frac{a_1}{1+a_1}+\frac{a_2}{1+a_2}-\frac{a_2a_4}{1+a_4}+\frac{a_5}{1+a_5}-\frac{a_6a_7}{1+a_7}+\frac{a_8}{1+a_8}-\cdots, \tag{3.4}$$

$$\frac{1}{1+a_1}-\frac{a_1a_2}{1+a_2}+\frac{a_3}{1+a_3}-\frac{a_4a_5}{1+a_5}+\frac{a_6}{1+a_6}-\frac{a_7a_8}{1+a_8}+\cdots \tag{3.5}$$

$$\frac{1}{1}+\frac{a_1}{1+a_1}-\frac{a_2a_3}{1+a_3}+\frac{a_4}{1+a_4}-\frac{a_5a_6}{1+a_6}+\frac{a_7}{1+a_7}-\cdots \tag{3.6}$$

have the sequences of approximants:

$$f_1, f_2, f_4, f_5, f_7, f_8, \cdots, f_{3p+1}, f_{3p+2}, \cdots,$$
$$f_0, f_2, f_3, f_5, f_6, f_8, \cdots, f_{3p}, f_{3p+2}, \cdots,$$
$$f_0, f_1, f_3, f_4, f_6, f_7, \cdots, f_{3p}, f_{3p+1}, \cdots,$$

respectively. This can be proved by mathematical induction with the aid of the recurrence formulas (1.1) (with $q=0$). The convergence of any two of these continued fractions implies the convergence of the third; and $f(a)$ converges only in case these three continued fractions converge. Now, (3.4) may be thrown into the form

$$1+\frac{1}{b_1}+\frac{1}{b_2}+\cdots,$$

and diverges if $\sum|b_p|$ converges, *i.e.*, (3.4) diverges if the series (3.1) converges. Likewise, (3.5) diverges if the series (3.2) converges and (3.6) diverges if the series (3.3) converges.

Example. If $a_p=-p(p+2)/(p+1)^2$, $p=1, 2, 3, \cdots$, $f(a)$ diverges.

THEOREM 3.3. *Suppose k is an integer greater than 1, $f(a)$ belongs to M_k, i and j are integers such that $0\leqq i$, $0<j-i<k$ and*

$$C_{ij}^{k}(p) = 1 - \frac{A_{k+j-i}^{kp+i} A_{k+i-j}^{kp+j}}{A_{k}^{kp+j} A_{k}^{kp+i}}, \qquad p = 0, 1, \cdots.$$

If the sequence $\{C_{ij}^{k}(p)\}_{p=0}^{\infty}$ *has a bounded infinite subsequence, then* $L_i = L_j$.

Proof. By (1.6), $[f_{kp+i}, f_{kp+j}, f_{k(p+1)+i}, f_{k(p+1)+j}] = C_{ij}^{k}(p)$ or

$$(f_{kp+i} - f_{kp+j})(f_{k(p+1)+i} - f_{k(p+1)+j}) = C_{ij}(p) \cdot (f_{kp+i} - f_{k(p+1)+i})(f_{kp+j} - f_{k(p+1)+j}).$$

Under the stated hypothesis, it follows that $L_i = L_j$.

4. Limaçons and cardioids. If the sequence of approximants of the continued fraction f' is some permutation of the sequence of approximants of $f(a)$, we may be able to deduce new convergence conditions for $f(a)$ by applying known conditions to f'. This is the idea used in [6] and further developed in [3].

We begin with the following theorem of [3]:

THEOREM 4.1. *If* $f'_{2p} = f_{2p+1}$ *and* $f'_{2p+1} = f_{2p}$, $p = 0, 1, \cdots$, *and*

$$(4.1) \qquad a'_1 = \frac{1 + a_2}{a_1}, \quad a'_{2p} = a_{2p} \text{ and } a'_{2p+1} = \frac{(1 + a_{2p})(1 + a_{2p+2})}{a_{2p+1}}, \quad p = 1, 2, \cdots$$

then the sequence of approximants of $1 - f(a')$ *is* $\{f'_p\}_{p=0}^{\infty}$.

Proof. By equations (1.7) and (1.6), if p is a nonnegative integer, $a'_{2p+2} = -[f'_{2p}, f'_{2p+1}, f'_{2p+2}, f'_{2p+3}] = -[f_{2p+1}, f_{2p}, f_{2p+3}, f_{2p+2}] = a_{2p+2}$; $a'_{2p+3} = -[f'_{2p+1}, f'_{2p+2}, f'_{2p+3}, f'_{2p+4}] = -[f_{2p}, f_{2p+3}, f_{2p+2}, f_{2p+5}] = (1+a_{2p+2})(1+a_{2p+4})/a_{2p+3}$; and $1 - (1+a'_1)^{-1} = f_3$ or $a'_1 = (1+a_2)/a_1$.

The known convergence theorem to be applied to $f(a')$ is [2]:

THEOREM 4.2. *If the sequence* a *satisfies the inequalities*

$$(4.2) \qquad \begin{aligned} &|1 + a_1| > |a_1|, \quad |1 + a_1 + a_2| > |a_2|, \\ &|1 + a_p + a_{p+1}| \geqq |a_p| + |a_{p+1}|, \qquad p = 2, 3, \cdots, \end{aligned}$$

then the sequences $\{f_{2p+1}\}$ *and* $\{f_{2p+2}\}$ *are absolutely convergent and* $f(a)$ *converges only in case the series* $\sum |b_p|$ *defined by* (2.2) *is divergent.*

Note. If [8] $|a_p| \leqq 1/4$, $p = 1, 2, \cdots$, then the inequalities (4.2) are satisfied and $\sum |b_p|$ diverges. Now, $|a'_p| \leqq 1/4$, $p = 1, 2, \cdots$, provided $|a_{2p}| \leqq 1/4$ and $|a_{2p-1}| \geqq 25/4$, $p = 1, 2, \cdots$. Hence, $f(a)$ converges if [3]:

$$(4.3) \qquad |a_{2p}| \leqq 1/4 \text{ and } |a_{2p-1}| \geqq 25/4, \qquad p = 1, 2, \cdots.$$

THEOREM 4.3. *Suppose, for each positive integer* p, a_{2p} *is a complex number distinct from* 0 *such that* $\operatorname{Re} a_{2p} > -1/2$, $c_p = 1 + a_{2p}$, $r_p = |a_{2p}/(1+a_{2p})|$, *so that* $0 < r_p < 1$, *and* a_{2p-1} *is a complex number distinct from* 0 *such that*

$$|a_1 + c_1| > |c_1|, \qquad |a_1 + 1| > r_1 |a_1|,$$

$$|a_{2p+1}+c_{p+1}| \geqq r_p|a_{2p+1}| + |c_{p+1}|, \quad |a_{2p+1}+c_p| \geqq r_{p+1}|a_{2p+1}| + |c_p|.$$

Then, the sequences $\{f_{2p+1}\}$ *and* $\{f_{2p+2}\}$ *converge absolutely and* $f(a)$ *converges only in case the series* $\sum|b_p|$ *defined by* (2.2) *is divergent.*

Proof. Under the hypothesis of the theorem, the sequence $\{a_p'\}$ defined by (4.1) satisfies the inequalities (4.2). Therefore, by Theorem 4.2, the sequences $\{f_{2p+1}\}$ and $\{f_{2p+2}\}$ are absolutely convergent; and $f(a')$ converges only if the series $\sum|b_p'|$ diverges, where $b_1'=1$ and $b_{p+1}'=1/a_p'b_p'$, $p=1, 2, \cdots$. Thus, $f(a)$ converges only if $\sum|b_p'|$ diverges. It remains to be shown that the series $\sum|b_p'|$ diverges if the series $\sum|b_p|$ defined by (2.2) diverges.

Suppose $\sum|b_p|$ diverges and consider two cases according as $\sum|a_{2p}|^{-1}$ converges or diverges. Since $|a_{2p}|^{-1}=|b_{2p}'b_{2p+1}'|$, we see that $\sum|b_p'|$ diverges in the second case. In the first case, there exists a positive number r such that $r_1r_2\cdots r_p\to r$ as $p\to\infty$. Since $|b_{2p+2}'|=|b_2'/b_3|\cdot|b_{2p+3}|\cdot(r_1r_2\cdots r_p)(r_2r_3\cdots r_{p+1})$ and $|b_{2p+3}'|=|b_3'/b_2|\cdot|b_{2p+2}|/(r_1r_2\cdots r_p)(r_2r_3\cdots r_{p+1})$, it follows that $\sum|b_p'|$ diverges in the first case.

This completes a proof of Theorem 4.3.

Note. The only restriction on a_{2p} in this theorem is: $a_{2p}\neq 0$, $\operatorname{Re} a_{2p} > -1/2$; a_1 is exterior to each of two circles depending upon a_2; a_{2p+1} is not interior to the outer loops of either of two limaçons depending on a_{2p} and a_{2p+2}. The polar equations of these limaçons are:

$$\rho=\frac{2|c|}{1-r^2}[r-\cos(\theta-\gamma)], \qquad c=|c|e^{i\gamma}, \qquad (c, r)=(c_{p+1}, r_p), (c_p, r_{p+1}).$$

These conditions are satisfied and $f(a)$ converges if (*cf.* (4.3))

$$|a_{2p}| \leqq \frac{1}{4} \quad \text{and} \quad |a_{2p-1}| \geqq \frac{15}{4}, \qquad p=1, 2, 3, \cdots. \tag{4.4}$$

Actually, $f(a)$ converges if $|a_{2p}|\leqq\frac{1}{4}$ and $|a_{2p-1}|\geqq 9/4$, $p=1, 2, 3, \cdots$, [1].

Let P denote the parabolic disc consisting of all points z such that $|z|-\operatorname{Re} z\leqq\frac{1}{2}$. If the terms of the sequence a belong to P, the inequalities (4.2) are satisfied and therefore $f(a)$ converges only if the series $\sum|b_p|$ defined by (2.2) diverges. This is the *parabola theorem* [4]. If W is a point set containing P as proper subset, there exists a sequence a whose terms belong to W such that $\sum|b_p|$ diverges and $f(a)$ diverges [5]. However, if the terms of a *after the first* belong to P and $\operatorname{Re} a_1>-1/2$, (4.2) holds and $f(a)$ converges only if $\sum|b_p|$ diverges.

THEOREM 4.4. *Suppose, for each positive integer* p,

$$|a_{2p}|-\operatorname{Re} a_{2p}\leqq\frac{1}{2},$$

$$\operatorname{Re}\left\{\frac{1+a_2}{|1+a_2|}\cdot\frac{1}{a_1}\right\} > -\frac{1}{2|1+a_2|},$$

and

$$\left|\frac{1}{a_{2p+1}}\right| - \operatorname{Re}\left\{\frac{(1+a_{2p})(1+a_{2p+2})}{|(1+a_{2p})(1+a_{2p+2})|}\cdot\frac{1}{a_{2p+1}}\right\} \leqq \frac{1}{2|(1+a_{2p})(1+a_{2p+2})|}.$$

Then, the sequences $\{f_{2p+1}\}$ *and* $\{f_{2p+2}\}$ *converge absolutely and* $f(a)$ *converges only in case the series* $\sum|b_p|$ *defined by* (2.2) *diverges.*

Proof. The hypothesis implies that the terms of the sequence $\{a_p'\}$ defined by (4.1), after the first, belong to the parabolic disc P and $\operatorname{Re} a_1' > -\frac{1}{2}$. Thus, the inequalities (4.2) are satisfied by $\{a_p'\}$ so that the proof of Theorem 4.3 applies here.

In this theorem, the region of a_{2p+1} is a *cardioid plus its exterior.* In particular, $f(a)$ converges if

$$|a_{2p}| \leqq \frac{1}{4} \quad \text{and} \quad |a_{2p-1}| \geqq \frac{9}{4}, \qquad p = 1, 2, 3, \cdots, \tag{4.5}$$

(*cf.* (4.3), (4.4) and [1]).

5. Other rearrangements. We conclude with three other examples of rearrangements and corresponding convergence theorems.

(1) If $f_{3p}' = f_{3p}$, $f_{3p+1}' = f_{3p+2}$ and $f_{3p+2}' = f_{3p+1}$, $p = 0, 1, \cdots$, and

$$a_{3p+1}' = -\frac{a_{3p-1}}{1+a_{3p}+a_{3p+1}}, \quad a_{3p+2}' = \frac{1}{a_{3p+2}}, \quad a_{3p+3}' = -\frac{a_{3p+3}}{1+a_{3p+3}+a_{3p+4}}, \tag{5.1}$$

$p = 0, 1, \cdots$ (with $a_0 = 0$), then $f(a')/(1+a_1)$ has the sequence of approximants $\{f_p'\}_{p=0}^{\infty}$.

(2) If $f_{4p}' = f_{4p}$, $f_{4p+1}' = f_{4p+1}$, $f_{4p+2}' = f_{4p+3}$ and $f_{4p+3}' = f_{4p+2}$, $p = 0, 1, \cdots$, and

$$\begin{aligned} a_{4p+1}' &= \frac{a_{4p+1}}{(1+a_{4p})(1+a_{4p+2})}, \quad a_{4p+2}' = -\frac{a_{4p+2}}{1+a_{4p+2}}, \\ a_{4p+3}' &= \frac{1}{a_{4p+3}} \quad \text{and} \quad a_{4p+4}' = -\frac{a_{4p+4}}{1+a_{4p+4}}, \qquad p = 0, 1, \cdots, (a_0 = 0), \end{aligned} \tag{5.2}$$

then $f(a')$ has the sequence of approximants $\{f_p'\}_{p=0}^{\infty}$.

(3) If $f_{5p}' = f_{5p}$, $f_{5p+1}' = f_{5p+1}$, $f_{5p+2}' = f_{5p+2}$, $f_{5p+3}' = f_{5p+4}$ and $f_{5p+4}' = f_{5p+3}$, $p = 0, 1, \cdots$, and

$$\begin{aligned} a_{5p+1}' &= \frac{a_{5p+1}}{1+a_{5p}}, \quad a_{5p+2}' = \frac{a_{5p+2}}{1+a_{5p+3}}, \quad a_{5p+3}' = -\frac{a_{5p+3}}{1+a_{5p+3}}, \\ a_{5p+4}' &= \frac{1}{a_{5p+4}}, \quad a_{5p+4}' = -\frac{a_{5p+5}}{1+a_{5p+5}}, \qquad p = 0, 1, \cdots, (a_0 = 0), \end{aligned} \tag{5.3}$$

then $f(a')$ has the sequence of approximants $\{f_p'\}_{p=0}^{\infty}$.

We omit proofs analogous to that of Theorem 4.1.

If we apply Theorem 4.2 to $f(a')$, where a' is defined by (5.1), we obtain

THEOREM 5.1. *If* $|a_1|<1$, $|1+a_1+a_2|>|1+a_1|$ *and, for each positive integer* p,

$$|a_{3p}|+|a_{3p+1}|<1,$$

$$\left|a_{3p+2}+\left(1+\frac{a_{3p+1}}{1+a_{3p}}\right)\right| \geqq \left|\frac{a_{3p+1}}{1+a_{3p}}\right|\cdot|a_{3p+2}|+\left|1+\frac{a_{3p+1}}{1+a_{3p}}\right| \text{ and}$$

$$\left|a_{3p+2}+\left(1+\frac{a_{3p+3}}{1+a_{3p+4}}\right)\right| \geqq \left|\frac{a_{3p+3}}{1+a_{3p+4}}\right|\cdot|a_{3p+2}|+\left|1+\frac{a_{3p+3}}{1+a_{3p+4}}\right|,$$

then $f(a)$ *converges only in case the series* $\sum|b_p'|$ *diverges, where, in terms of* (5.1), $b_1'=1$ *and* $b_{p+1}'=1/b_p'a_p'$, $p=1, 2, 3, \cdots$.

In this case the boundaries of the regions are made up of arcs of circles and limaçons.

In the next two theorems, $1/P$ denotes the set to which z belongs only if $1/z$ belongs to the parabolic disc P defined in Section 4 and is a cardioid plus its exterior. The polar equation of this cardioid is $\rho=2(1-\cos\theta)$. Also, $-1-1/P$ denotes the set to which w belongs only if there is a point z in P such that $w=-1-1/z$ and $1/[-1-1/P]$ is the inversion in the unit circle of the region $-1-1/P$ and is a subset of the unit circular disc having -1 and $\frac{1}{3}$ on its boundary.

We apply the parabola theorem to $f(a')$, where a' is defined by (5.2), and obtain:

THEOREM 5.2. *If, for each nonnegative integer* p, a_{4p+3} *belongs to* $1/P$, a_{2p+2} *to* $1/[-1-1/P]$ *and* (*with* $a_0=0$)

$$|a_{4p+1}|-\operatorname{Re}\left\{\frac{|(1+a_{4p})(1+a_{4p+2})|}{(1+a_{4p})(1+a_{4p+2})}\cdot a_{4p-1}\right\} \leqq \frac{|(1+a_{4p})(1+a_{4p+2})|}{2},$$

then $f(a)$ *converges only in case the series* $\sum|b_p'|$ *diverges where, in terms of* a' *defined by* (5.2), $b_1'=1$ *and* $b_{p+1}'=1/b_p'a_p'$, $p=1, 2, \cdots$.

Finally, we apply the parabola theorem to $f(a')$, where a' is defined by (5.3), and get:

THEOREM 5.3. *If, for each nonnegative integer* p, a_{5p+4} *belongs to* $1/P$, a_{5p+3} *to* $1/[-1-1/P]$ *and, with* $a_0=0$,

$$|a_{5p+1}|-\operatorname{Re}\left\{\frac{|1+a_{5p}|}{1+a_{5p}}\cdot a_{5p+1}\right\} \leqq \frac{|1+a_{5p}|}{2}$$

and

$$|a_{5p+2}|-\operatorname{Re}\left\{\frac{|1+a_{5p+3}|}{1+a_{5p+3}}\cdot a_{5p+2}\right\} \leqq \frac{|1+a_{5p+3}|}{2},$$

then $f(a)$ converges only in case the series $\sum|b_p'|$ diverges where, in terms of (5.3), $b_1'=1$ *and* $b_{p+1}'=1/b_p'a_p'$, $p=1, 2, \cdots$.

Bibliography

1. Copp, George. Some convergence regions for continued fractions, Dissertation, The University of Texas, 1950.

2. Lane, R. E. and Wall, H. S. Continued fractions with absolutely convergent even and odd parts, Trans. Amer. Math. Soc., vol. 67, 1949, pp. 368–380.

3. Leighton, W. and Wall, H. S. On the transformation and convergence of continued fractions, Amer. J. of Math., vol. 58, 1936, pp. 267–281.

4. Scott, W. T. and Wall, H. S. A convergence theorem for continued fractions, Trans. Amer. Math. Soc., vol. 47, 1940, pp. 155–172.

5. Wall, B. M. Some convergence sets for a continued fraction, Dissertation, The University of Texas, 1956.

6. Wall, H. S. On continued fractions and cross-ratio groups of Cremona transformations, Bull. Amer. Math. Soc., vol. 40, 1934, pp. 587–592.

7. ———. Partially bounded continued fractions, Proc. Amer. Math. Soc., vol. 7, 1956, pp. 1090–1093.

8. Worpitzky, J. Untersuchungen über die Entwickelung der monodronen und monogenen Funktionen durch Kettenbrüche, Friedrichs-Gymnasium und Realschule, Jahresbericht, Berlin, 1865, pp. 3–39.

SENSE AND ORIENTATION ON THE DISK

G. T. WHYBURN, University of Virginia

We shall be concerned with a simple and elementary approach to some fundamental results having to do with orientation and sense agreement for closed curves in a plane or on a disk. These theorems, although of basic and crucial importance in a really complete treatment of topological and mapping results in this setting, are frequently passed over in the early stages and treated, if at all, only after quite elaborate or sophisticated techniques have been developed. Thus they occur as special cases or as corollaries to deep theorems whose statements and meaning may be concealed from the beginner by their involvement in elaborate concepts and notation.

The critical issue in this topic is to define *sense* on a simple closed curve on a disk and *agreement in sense* for two such curves J_1 and J_2 so that each curve has exactly two senses, and then to show that if a given sense on J_1 agrees with say the "positive" sense on J_2 then it *cannot* also agree with the "negative" sense on J_2. All the results to be proven are well known, of course. The ones selected are of interest since they may be used to prove such results as the invariance of the

property of orientability for a 2-manifold, for example—both the invariance under subdivision and the topological invariance—making little use of subdivision techniques. A discussion of this topic, as well as a number of the theorems in the present paper, may be found treated by subdivision modification methods in Kerekjarto's book on topology.

DEFINITION. *The ordering* (a, b, c) *of three points* a, b, c *on a circle* C *of the complex plane is a positive ordering provided that when* C *plus its interior is regarded as a simplex with vertices* a, b, c *then* (a, b, c) *is a positive orientation of this simplex in the complex plane. Equivalently,* (a, b, c) *is a positive ordering of* a, b, c *provided that, assuming* C *has center at the origin and radius* r*, if* $g(\theta)$ *is the mapping* $re^{i\theta}$ *of the interval* $(0, 2\pi)$ *onto* C *and if* $\theta_1<\theta_2<\theta_3$ *are the least three values on* $(0, 2\pi)$ *in* $g^{-1}(a+b+c)$*, then* $[g(\theta_1), g(\theta_2), g(\theta_3)]$ *is an even permutation of the ordering* (a, b, c).

THEOREM 1. *Given concentric circles* C *and* C' *of radii* r *and* r', $r'<r$, *and three disjoint simple arcs* aa', bb', cc' *with* $a+b+c\subset C$, $a'+b'+c'\subset C'$ *and each lying except for its end points in the annular region between* C *and* C'. *Then if* (a, b, c) *is a positive ordering of the points* a, b *and* c, (a', b', c') *is a positive ordering of the points* a', b' *and* c'.

Proof. Suppose on the contrary that for some choice of the arcs aa', bb', cc' it were true that the ordering (a', b', c') is negative (*i.e.*, not positive). Then without loss of generality we may assume† that $r'=\frac{3}{4}r$. Further we may assume each of the arcs is of finite length so that the sum s of their lengths is finite. Also if σ is the g.l.b. of the length sum s for all possible sets of 3 arcs aa', bb', cc' as above with (a, b, c) a positive ordering but (a', b', c') a negative ordering, we may suppose our particular choice of arcs made so that their length sum s satisfies $s<\sigma+\frac{1}{4}r'$.

At least one of the three arcs ab, bc, ca of C is of length $\geqq 2\pi r/3$. We may suppose this true of the arc ca. Let a_1 be the midpoint of ca. Then the radius a_1o of C cannot intersect $aa'+bb'+cc'=T$. For suppose it does. Then let p be the first point of T on a_1o in the order a_1, o. Since the open arc bb' contains points inside the simple closed curve $J=abc$ (of C)$+aa'+cc'+a'b'c'$ (of C') but does not intersect J, it therefore lies entirely inside J. Accordingly, as a_1p-p is entirely outside J it follows that p cannot belong to the arc bb'. Thus p lies on

† To see this, note that we can apply the preliminary radial stretching transformation $(\rho, \theta)\rightarrow(\rho^*, \theta^*)$ given by

$$\rho^* = \frac{3r}{4r'}\rho, \qquad \theta^* = \theta, \text{ for } \rho \leqq r'$$

$$\rho^* = \frac{r\rho + 3r^2 - 4rr'}{4(r-r')}, \qquad \theta^* = \theta, \text{ for } \rho \geqq r',$$

which will leave the origin and all points of C fixed and will not alter order relations of points on C' but maps C' into the circle with center O and radius $\frac{3}{4}r$.

either aa' or cc'. If $p \in aa'$, we replace the arc ap in aa' by the arc a_1p of a_1o; and this clearly diminishes the length sum s by more than $\frac{1}{4}r'$ because the length a_1p is $<\frac{1}{4}r$, whereas the length of ap (in aa') is $\geqq |a-a_1| - |a_1-p| \geqq r-\frac{1}{4}r = \frac{3}{4}r$. This makes the length sum of the new arcs a_1a', bb', cc' less than σ, which is impossible because (a_1, b, c) is still a positive ordering of the points a_1, b, c on C. Similarly if $p \in cc'$ we replace the arc cp of cc' by the arc a_1p of a_1o and reason as before; (a_1, a, b) is still a positive ordering of a_1, a, b so that the same contradiction is obtained. Hence T does not intersect the radius a_1o of C.

Now let c_1o be a radius of C with c_1 on the arc a_1c of C and so close to a_1o that no point of T is in the sector of C bounded by $a_1o + c_1o +$ arc a_1c_1 of C. Let c_1c_1' and a_1a_1' be the segments of c_1o and a_1o respectively such that c_1' and a_1' are on C'. Consider the simple closed curve $K =$ arc a_1abcc_1 (of C) $+ c_1c_1' +$ arc $c_1'a_1'$ of C' containing a', b' and $c' + a_1'a_1$. Since b and b' separate a and c on K and since the arcs aa', bb', cc' lie in $K+$ its interior but do not intersect, it follows that b and b' also separate a' and c' on K and that a and a' lie on one of the arcs of K from b to b' and c and c' on the other. Accordingly we must have the order $a_1'a'b'c'c_1'$ on the positively oriented arc which K has in common with C'. Thus (a', b', c') is a positive ordering of the points a', b', c' contrary to our supposition.

DEFINITION. *Let A be a 2-cell with edge J and let J' be a simple closed curve on A lying interior to J. If a, b, c and a', b', c' are triples of distinct points on J and J' respectively, the sense $a'b'c'$ on J' given by the ordering (a', b', c') is said to agree with the sense abc on J given by the ordering (a, b, c) provided there exist disjoint simple arcs aa', bb', cc' lying except for their ends in the annular region of A bounded by J and J'.*

THEOREM 2. *If the sense abc on J agrees with the sense $a'b'c'$ on J', it cannot agree with the sense $a'c'b'$. Thus abc agrees with the sense given by any even permutation of the ordering (a', b', c') but with no odd one.*

Proof. Let G denote the graph on A consisting of J, J' and the arcs aa', bb', cc' and let h be a homeomorphism mapping G into the complex plane so that J and J' map onto the circles $|z| = 2$ and $|z| = 1$ respectively with a, b, c going into 2, $2e^{2\pi i/3}$, $2e^{4\pi i/3}$ respectively and a', b', c' into 1, $e^{2\pi i/3}$, $e^{4\pi i/3}$ respectively and with the arcs aa', bb', cc' mapping onto the linear segments joining the images of their respective end points. Since each region on A complementary to G is bounded by a simple closed curve of G whose image bounds a region in $|z| \leqq 2$ complementary to $h(G)$, h can be extended to a homeomorphism of A onto $|z| \leqq 2$.

Now if, contrary to our theorem, it were possible to find disjoint arcs aa', bc', cb' on A, then $h(aa')$, $h(bc')$, $h(cb')$ would be disjoint arcs in $1 \leqq |z| \leqq 2$; and since $(2, e^{2\pi i/3}, e^{4\pi i/3}) = [h(a), h(b), h(c)]$ is a positive ordering on the circle $|z| = 2$, it would follow by Theorem 1 that $[h(a'), h(c'), h(b')]$ has to be a positive ordering on $|z| = 1$. This is impossible because $h(a') = 1$, $h(c') = e^{4\pi i/3}$, $h(b') = e^{2\pi i/3}$.

COROLLARY 1. *If A is a 2-cell with edge J and $h(A)=B$ is any homeomorphism, then if the sense $a'b'c'$ on a simple closed curve J' interior to J on A agrees with the sense abc on J, the sense $h(a')h(b')h(c')$ on $h(J')$ must agree with $h(a)h(b)h(c)$ on $h(J)$.*

COROLLARY 2. *Any homeomorphism of an annular ring onto itself which maps one edge onto itself in sense-preserving fashion must likewise map the other edge on itself in a sense-preserving fashion.*

COROLLARY 3. *Any homeomorphism $h(A)=A$ of a 2-cell onto itself which preserves sense on the edge J of A must likewise preserve sense for all simple closed curves interior to A, i.e., if J' is any such curve interior to A and $a'b'c'$ is a sense on J' agreeing with the sense abc on J then the sense $h(a')h(b')h(c')$ on $h(J)$ will necessarily also agree with the sense abc on J.*

Note. By a *curved* triangle (or 2-simplex) is meant a set homeomorphic with an ordinary straight triangle in which sides and vertices are designated on its edge. An orientation of such a triangle is determined by an ordering of its vertices as discussed above.

THEOREM 3. *If Δ_1 and Δ_2 are curved triangles with vertices x, y, z and x, z, w respectively lying inside a 2-cell Δ and having a side xz and only this side in common and if (x, y, z) and (x, z, w) are orientations of Δ_1 and Δ_2 respectively each of which agrees with the orientation (a', b', c') of Δ, then these orientations must agree with each other in that their ordering on the common side xz is opposite.*

Proof. Let x, y and z be joined to points of the edge J of Δ by disjoint simple arcs xa, yb and zc lying except for their ends in the annular region on Δ between J and the simple closed curve $xy+yz+zw+wx$. Then since the ordering xyz of the boundary of Δ_1 agrees with the ordering abc on J in sense, we can choose a, b and c as vertices of Δ and still have the given orientations of Δ_1 and Δ_2 agree with Δ. Now the arc $axzc$ divides the interior of Δ into two regions R_w and R_y containing w and y respectively. Also R_w+c contains a simple arc wc. Thus ax, $by+$side yz of Δ_1 and cw are disjoint arcs in the annulus bounded by J and the edge of Δ_2. Accordingly, the sense given by ordering (x, z, w) on the edge of Δ_2 agrees with the sense abc on J. Since xzw and $yzx(=xyz)$ are opposite on the common side xz, our result is established.

Dedication. This paper is dedicated to Lester R. Ford on the occasion of his seventieth birthday. In the past 30 years, throughout which it has been the author's privilege to know and be associated with Dr. Ford in various mathematical and personal endeavors, his stimulating personality, his keen and zestful perception in all phases of mathematics, and above all his unfailing kindly interest in his fellow man have greatly enriched the author's life and added much to his mathematical experience.